GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 11:56:31 ; Search time 159.884 Seconds
(without alignments)
1852.232 Million cell updates/sec

Title: US-10-775-180-447
Perfect score: 3568
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....TCFAEEGKKLVAASQAALGL 674

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3568	100.0	674	7	ADF16193	Adf16193 Human alb
2	3568	100.0	674	7	ADH21650	Adh21650 Human alb
3	3568	100.0	674	9	ADW45202	Adw45202 K. lactis
4	3568	100.0	915	9	ADW45204	Adw45204 K. lactis
5	3444.5	96.5	669	7	ADF16144	Adf16144 Human alb
6	3444.5	96.5	669	7	ADH21622	Adh21622 Human alb
7	3444	96.5	730	7	ADF16525	Adf16525 Human alb
8	3444	96.5	730	7	ADH21813	Adh21813 Human alb
9	3438.5	96.4	669	7	ADF16150	Adf16150 Human alb
10	3438.5	96.4	669	7	ADH21628	Adh21628 Human alb
11	3438	96.4	730	7	ADF16527	Adf16527 Human alb
12	3438	96.4	730	7	ADH21815	Adh21815 Human alb
13	3432.5	96.2	669	7	ADF16149	Adf16149 Human alb
14	3432.5	96.2	669	7	ADF16148	Adf16148 Human alb
15	3432.5	96.2	669	7	ADF16145	Adf16145 Human alb
16	3432.5	96.2	669	7	ADF16146	Adf16146 Human alb
17	3432.5	96.2	669	7	ADH21624	Adh21624 Human alb
18	3432.5	96.2	669	7	ADH21626	Adh21626 Human alb
19	3432.5	96.2	669	7	ADH21623	Adh21623 Human alb
20	3432.5	96.2	669	7	ADH21627	Adh21627 Human alb
21	3427	96.0	668	7	ADF16524	Adf16524 Human alb
22	3427	96.0	668	7	ADH21812	Adh21812 Human alb
23	3422	95.9	662	7	ADF16526	Adf16526 Human alb
24	3422	95.9	662	7	ADH21814	Adh21814 Human alb
25	3421	95.9	668	7	ADF16528	Adf16528 Human alb
26	3421	95.9	668	7	ADH21816	Adh21816 Human alb
27	3420.5	95.9	664	7	ADF16510	Adf16510 Human alb
28	3420.5	95.9	664	7	ADH21801	Adh21801 Human alb
29	3418.5	95.8	663	7	ADF16512	Adf16512 Human alb
30	3418.5	95.8	663	7	ADH21803	Adh21803 Human alb
31	3416	95.7	662	7	ADF16529	Adf16529 Human alb
32	3416	95.7	662	7	ADH21817	Adh21817 Human alb
33	3414.5	95.7	664	7	ADF16511	Adf16511 Human alb
34	3414.5	95.7	664	7	ADH21802	Adh21802 Human alb
35	3413.5	95.7	667	7	ADF16147	Adf16147 Human alb
36	3413.5	95.7	667	7	ADH21625	Adh21625 Human alb
37	3412.5	95.6	663	7	ADF16513	Adf16513 Human alb
38	3412.5	95.6	663	7	ADH21804	Adh21804 Human alb
39	3401	95.3	658	9	ADW45206	Adw45206 K. lactis
40	3395	95.2	654	9	ADW45215	Adw45215 K. lactis
41	3395	95.2	656	9	ADW45221	Adw45221 K. lactis
42	3393	95.1	650	9	ADW45205	Adw45205 K. lactis
43	3392.5	95.1	655	9	ADW45216	Adw45216 K. lactis
44	3392.5	95.1	657	9	ADW45212	Adw45212 K. lactis
45	3392.5	95.1	657	9	ADW45299	Adw45299 Human fus

ALIGNMENTS

RESULT 1
ADF16193
ID ADF16193 standard; protein; 674 AA.
XX
AC ADF16193;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human albumin therapeutic fusion protein SeqID1280.
XX
KW albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human.
XX
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003060071-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040891.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.

10/775,204
2005-0186664

PA (PRIN-) PRINCIPIA PHARM CORP.
XX
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598517/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Example 4; SEQ ID NO 1280; 24pp; English.
XX
CC This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is the amino acid sequence of a
CC novel full-length human albumin therapeutic fusion protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ Sequence 674 AA;

Query Match 100.0%; Score 3568; DB 7; Length 674;
Best Local Similarity 100.0%; Pred. No. 3.8e-293;
Matches 674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        661 GKKLVAASQAALGL 674
```

RESULT 2
ADH21650
ID ADH21650 standard; protein; 674 AA.
XX
AC ADH21650;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human albumin/GLP-1(7-36(A8G))x2 fusion protein, SEQ ID NO:447.
XX
KW Fusion protein; human serum albumin; HSA; therapeutic protein;
KW shelf-life; in vitro biological activity; in vivo biological activity;
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW anorectic; ophthalmological; gene therapy.
XX
OS Synthetic.
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003059934-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040892.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598501/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Disclosure; SEQ ID NO 447; 1086pp; English.
XX
CC The invention relates to fusion proteins comprising human serum albumin
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC antibody or peptide or their variants or fragments. The therapeutic
CC protein may be fused to the N-terminus, the C-terminus or both termini of
CC albumin via a linker. The albumin component of the fusion proteins
CC prolongs the shelf-life and the in vitro and vivo biological activity of
CC the proteins compared with those of the corresponding therapeutic
CC proteins on their own. The invention also relates to nucleic acids
CC encoding albumin fusion proteins, vectors and host cells comprising an
CC albumin fusion protein nucleic acid, compositions and kits comprising an

10/775 180

CC albumin fusion protein, the method of extending the shelf-life of a
CC therapeutic protein by fusion with albumin, and the treatment of disease
CC using an albumin fusion protein. The albumin fusion proteins may be used
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC related conditions. Specifically the albumin fusion proteins may be used
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC (especially neuropathy), retinopathy, cardiovascular disorders
CC (especially heart disease, renal disorders and obesity. The proteins may
CC also be used in a method of maintaining a basal glucose level in a
CC patient and in a method for losing weight. The present sequence is
CC related to the invention.
XX
SQ Sequence 674 AA;

Query Match 100.0%; Score 3568; DB 7; Length 674;
Best Local Similarity 100.0%; Pred. No. 3.8e-293;
Matches 674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        661 GKKLVAASQAALGL 674
```

RESULT 3
ADW45202
ID ADW45202 standard; protein; 674 AA.
XX
AC ADW45202;
XX
DT 07-APR-2005 (first entry)
XX
DE K. lactis killer toxin-GLP1-human serum albumin fusion protein - SEQ 206.
XX
KW fusion protein; anti-HIV; gastrointestinal-gen.; antidiabetic; anorectic;
KW nephrotropic; cardiant; cytostatic; neuroprotective; immunosuppressive;
KW immune disorder; hematological disease; hyperproliferative disorder;
KW renal disease; cardiovascular disease; cardiovascular-gen.;
KW respiratory disorder; angiogenesis disorder; neurological disease;
KW wound healing; vulnerary; endocrine disease; reproductive disorder;
KW gynecological; infectious disease; antimicrobial;
KW gastrointestinal disease; gene therapy; toxin; HSA; albumin;
KW glucagon-like peptide 1; GLP1.
XX
OS Homo sapiens.
OS Kluyveromyces lactis.
OS Chimeric.
XX
PN WO2005003296-A2.
XX
PD 13-JAN-2005.
XX
PF 20-JAN-2004; 2004WO-US001369.
XX
PR 22-JAN-2003; 2003US-0441305P.
PR 11-MAR-2003; 2003US-0453201P.
PR 02-MAY-2003; 2003US-0467222P.
PR 23-MAY-2003; 2003US-0472816P.
PR 06-JUN-2003; 2003US-0476267P.
PR 24-SEP-2003; 2003US-0505172P.
PR 30-SEP-2003; 2003US-0506746P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Haseltine WA, Rosen CA;
XX
DR WPI; 2005-091786/10.
XX
PT New albumin fusion protein for diagnosing, treating or preventing
PT diseases such as HIV/AIDS, diabetes, obesity, heart disease or immune
PT disorders comprises a therapeutic protein (e.g. CD4M33, GLP-2 or PACAP-
PT 27) and an albumin.
XX
PS Example 13; SEQ ID NO 206; 884pp; English.
XX
CC The invention relates to a novel albumin fusion protein comprising a
CC therapeutic protein as listed in the specification in Table 1 and an
CC albumin comprising a sequence of SEQ ID NO: 1, or a fragment or variant
CC of SEQ ID NO: 1, where the fragment or variant has albumin activity and
CC where the albumin activity is the ability to prolong the shelf life of
CC the therapeutic protein compared to the shelf-life of the therapeutic
CC protein in an unfused state. Human serum albumin (HSA, HA) is responsible
CC for a significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion
CC protein of the invention demonstrates anti-HIV, gastrointestinal-gen.,
CC antidiabetic, anorectic, cardiant and immunosuppressive activities. The
CC fusion protein may be useful for diagnosing, treating, preventing or
CC ameliorating diseases, such as immune disorders, blood disorders,
CC hyperproliferative disorders, renal disorders, cardiovascular disorders,
CC respiratory disorders, angiogenesis-related disorders, neurological
CC disorders, wound healing disorders, endocrine disorders, reproductive
CC disorders, infectious disorders and gastrointestinal disorders, possibly
CC with the use of gene therapy techniques. The current sequence is that of
CC the Kluyveromyces lactis killer toxin-GLP1-human serum albumin fusion
CC protein - SEQ 206 of the invention.
XX
SQ Sequence 674 AA;

Query Match 100.0%; Score 3568; DB 9; Length 674;
Best Local Similarity 100.0%; Pred. No. 3.8e-293;
Matches 674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
```

11/175,690
2006/0014254

```
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
Db        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
Db        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
Db        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
Db        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
Db        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
Db        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
Db        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
Db        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
Db        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Qy        661 GKKLVAASQAALGL 674
Db        661 GKKLVAASQAALGL 674
```

RESULT 4
ADW45204
ID ADW45204 standard; protein; 915 AA.
XX
AC ADW45204;
XX
DT 07-APR-2005 (first entry)
XX
DE K. lactis killer toxin-GLP1-HSA-GFP tag fusion protein - SEQ ID 208.
XX
KW fusion protein; anti-HIV; gastrointestinal-gen.; antidiabetic; anorectic;
KW nephrotropic; cardiant; cytostatic; neuroprotective; immunosuppressive;
KW immune disorder; hematological disease; hyperproliferative disorder;
KW renal disease; cardiovascular disease; cardiovascular-gen.;
KW respiratory disorder; angiogenesis disorder; neurological disease;
KW wound healing; vulnerary; endocrine disease; reproductive disorder;
KW gynecological; infectious disease; antimicrobial;
KW gastrointestinal disease; gene therapy; toxin; HSA; albumin;
KW glucagon-like peptide 1; GLP1.
XX
OS Homo sapiens.
OS Kluyveromyces lactis.
OS Chimeric.
XX
PN WO2005003296-A2.
XX
PD 13-JAN-2005.
XX
PF 20-JAN-2004; 2004WO-US001369.
XX
PR 22-JAN-2003; 2003US-0441305P.
PR 11-MAR-2003; 2003US-0453201P.
PR 02-MAY-2003; 2003US-0467222P.
PR 23-MAY-2003; 2003US-0472816P.
PR 06-JUN-2003; 2003US-0476267P.
PR 24-SEP-2003; 2003US-0505172P.
PR 30-SEP-2003; 2003US-0506746P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Haseltine WA, Rosen CA;
XX
DR WPI; 2005-091786/10.
XX
PT New albumin fusion protein for diagnosing, treating or preventing
PT diseases such as HIV/AIDS, diabetes, obesity, heart disease or immune
PT disorders comprises a therapeutic protein (e.g. CD4M33, GLP-2 or PACAP-
PT 27) and an albumin.
XX
PS Example 13; SEQ ID NO 208; 884pp; English.
XX
CC The invention relates to a novel albumin fusion protein comprising a
CC therapeutic protein as listed in the specification in Table 1 and an
CC albumin comprising a sequence of SEQ ID NO: 1, or a fragment or variant
CC of SEQ ID NO: 1, where the fragment or variant has albumin activity and
CC where the albumin activity is the ability to prolong the shelf life of
CC the therapeutic protein compared to the shelf-life of the therapeutic
CC protein in an unfused state. Human serum albumin (HSA, HA) is responsible
CC for a significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion
CC protein of the invention demonstrates anti-HIV, gastrointestinal-gen.,
CC antidiabetic, anorectic, cardiant and immunosuppressive activities. The
CC fusion protein may be useful for diagnosing, treating, preventing or
CC ameliorating diseases, such as immune disorders, blood disorders,
CC hyperproliferative disorders, renal disorders, cardiovascular disorders,
CC respiratory disorders, angiogenesis-related disorders, neurological
CC disorders, wound healing disorders, endocrine disorders, reproductive
CC disorders, infectious disorders and gastrointestinal disorders, possibly
CC with the use of gene therapy techniques. The current sequence is that of
CC the Kluyveromyces lactis killer toxin-GLP1-HSA-GFP tag fusion protein -
CC SEQ ID 208 of the invention.
XX
SQ Sequence 915 AA;

Query Match 100.0%; Score 3568; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 5.7e-293;
Matches 674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
Db        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
Db        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
Db        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
```

```
Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        661 GKKLVAASQAALGL 674
```

```
RESULT 5
ADF16144
ID   ADF16144 standard; protein; 669 AA.
XX
AC   ADF16144;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1231.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1231; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 669 AA;

  Query Match             96.5%;  Score 3444.5;  DB 7;  Length 669;
  Best Local Similarity   97.3%;  Pred. No. 1.1e-282;
  Matches  654;  Conservative    4;  Mismatches    5;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :          ||||||||||||||||||||||||||||||||||||
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 LKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 LKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669


RESULT 6
ADH21622
ID   ADH21622 standard; protein; 669 AA.
XX
AC   ADH21622;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:419.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Disclosure; SEQ ID NO 419; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 669 AA;

  Query Match             96.5%;  Score 3444.5;  DB 7;  Length 669;
  Best Local Similarity   97.3%;  Pred. No. 1.1e-282;
  Matches  654;  Conservative    4;  Mismatches    5;  Indels    9;  Gaps    1;

Qy          3 IFYIPLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :          ||||||||||||||||||||||||||||||||||||
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669


RESULT 7
ADF16525
ID   ADF16525 standard; protein; 730 AA.
```

```
XX
AC   ADF16525;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1622.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1622; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 730 AA;

  Query Match             96.5%;  Score 3444;  DB 7;  Length 730;
  Best Local Similarity   97.8%;  Pred. No. 1.4e-282;
  Matches  652;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 183

Qy        128 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 243

Qy        188 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 303

Qy        248 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 363

Qy        308 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 423

Qy        368 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 483

Qy        428 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 543

Qy        488 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 603

Qy        548 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 663

Qy        608 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 723

Qy        668 SQAALGL 674
              |||||||
Db        724 SQAALGL 730


RESULT 8
ADH21813
ID   ADH21813 standard; protein; 730 AA.
XX
AC   ADH21813;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:610.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
```

XX
PN WO2003059934-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040892.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598501/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Disclosure; SEQ ID NO 610; 1086pp; English.
XX
CC The invention relates to fusion proteins comprising human serum albumin
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC antibody or peptide or their variants or fragments. The therapeutic
CC protein may be fused to the N-terminus, the C-terminus or both termini of
CC albumin via a linker. The albumin component of the fusion proteins
CC prolongs the shelf-life and the in vitro and vivo biological activity of
CC the proteins compared with those of the corresponding therapeutic
CC proteins on their own. The invention also relates to nucleic acids
CC encoding albumin fusion proteins, vectors and host cells comprising an
CC albumin fusion protein nucleic acid, compositions and kits comprising an
CC albumin fusion protein, the method of extending the shelf-life of a
CC therapeutic protein by fusion with albumin, and the treatment of disease
CC using an albumin fusion protein. The albumin fusion proteins may be used
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC related conditions. Specifically the albumin fusion proteins may be used
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC (especially neuropathy), retinopathy, cardiovascular disorders
CC (especially heart disease, renal disorders and obesity. The proteins may
CC also be used in a method of maintaining a basal glucose level in a
CC patient and in a method for losing weight. The present sequence is
CC related to the invention.
XX
SQ Sequence 730 AA;

Query Match 96.5%; Score 3444; DB 7; Length 730;
Best Local Similarity 97.8%; Pred. No. 1.4e-282;
Matches 652; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : : :       ||||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 183

Qy        128 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 243

Qy        188 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 303

Qy        248 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 363

Qy        308 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 423

Qy        368 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 483

Qy        428 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 543

Qy        488 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 603

Qy        548 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 663

Qy        608 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 723

Qy        668 SQAALGL 674
              |||||||
Db        724 SQAALGL 730
```

RESULT 9
ADF16150
ID ADF16150 standard; protein; 669 AA.
XX
AC ADF16150;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human albumin therapeutic fusion protein SeqID1237.
XX
KW albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human.
XX
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003060071-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040891.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.

PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
PA (PRIN-) PRINCIPIA PHARM CORP.
XX
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598517/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Example 4; SEQ ID NO 1237; 24pp; English.
XX
CC This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is the amino acid sequence of a
CC novel full-length human albumin therapeutic fusion protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ Sequence 669 AA;

Query Match 96.4%; Score 3438.5; DB 7; Length 669;
Best Local Similarity 97.2%; Pred. No. 3.5e-282;
Matches 653; Conservative 4; Mismatches 6; Indels 9; Gaps 1;

```
Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||||||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669
```

RESULT 10
ADH21628
ID ADH21628 standard; protein; 669 AA.
XX
AC ADH21628;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:425.
XX
KW Fusion protein; human serum albumin; HSA; therapeutic protein;
KW shelf-life; in vitro biological activity; in vivo biological activity;
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW anorectic; ophthalmological; gene therapy.
XX
OS Synthetic.
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003059934-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040892.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598501/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX

PS Disclosure; SEQ ID NO 425; 1086pp; English.
XX
CC The invention relates to fusion proteins comprising human serum albumin
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC antibody or peptide or their variants or fragments. The therapeutic
CC protein may be fused to the N-terminus, the C-terminus or both termini of
CC albumin via a linker. The albumin component of the fusion proteins
CC prolongs the shelf-life and the in vitro and vivo biological activity of
CC the proteins compared with those of the corresponding therapeutic
CC proteins on their own. The invention also relates to nucleic acids
CC encoding albumin fusion proteins, vectors and host cells comprising an
CC albumin fusion protein nucleic acid, compositions and kits comprising an
CC albumin fusion protein, the method of extending the shelf-life of a
CC therapeutic protein by fusion with albumin, and the treatment of disease
CC using an albumin fusion protein. The albumin fusion proteins may be used
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC related conditions. Specifically the albumin fusion proteins may be used
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC (especially neuropathy), retinopathy, cardiovascular disorders
CC (especially heart disease, renal disorders and obesity. The proteins may
CC also be used in a method of maintaining a basal glucose level in a
CC patient and in a method for losing weight. The present sequence is
CC related to the invention.
XX
SQ Sequence 669 AA;

Query Match 96.4%; Score 3438.5; DB 7; Length 669;
Best Local Similarity 97.2%; Pred. No. 3.5e-282;
Matches 653; Conservative 4; Mismatches 6; Indels 9; Gaps 1;

```
Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :          ||||||||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657
```

```
Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669
```

RESULT 11
ADF16527
ID ADF16527 standard; protein; 730 AA.
XX
AC ADF16527;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human albumin therapeutic fusion protein SeqID1624.
XX
KW albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human.
XX
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003060071-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040891.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
PA (PRIN-) PRINCIPIA PHARM CORP.
XX
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598517/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Example 4; SEQ ID NO 1624; 24pp; English.
XX
CC This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating

```
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 730 AA;

  Query Match             96.4%;  Score 3438;  DB 7;  Length 730;
  Best Local Similarity   97.6%;  Pred. No. 4.4e-282;
  Matches  651;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :        |||||||||||||||||||||||||||||||||||| ||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 183

Qy        128 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 243

Qy        188 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 303

Qy        248 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 363

Qy        308 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 423

Qy        368 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 483

Qy        428 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 543

Qy        488 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 603

Qy        548 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 663

Qy        608 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 723

Qy        668 SQAALGL 674
              |||||||
Db        724 SQAALGL 730


RESULT 12
ADH21815
ID   ADH21815 standard; protein; 730 AA.
XX
AC   ADH21815;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:612.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Disclosure; SEQ ID NO 612; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 730 AA;

  Query Match             96.4%;  Score 3438;  DB 7;  Length 730;
  Best Local Similarity   97.6%;  Pred. No. 4.4e-282;
  Matches  651;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :        |||||||||||||||||||||||||||||||||||| ||||||
```

```
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 183

Qy        128 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 243

Qy        188 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 303

Qy        248 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 363

Qy        308 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 423

Qy        368 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 483

Qy        428 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 543

Qy        488 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 603

Qy        548 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 663

Qy        608 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 723

Qy        668 SQAALGL 674
              |||||||
Db        724 SQAALGL 730
```

```
RESULT 13
ADF16149
ID   ADF16149 standard; protein; 669 AA.
XX
AC   ADF16149;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1236.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1236; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 669 AA;
```

```
  Query Match             96.2%;  Score 3432.5;  DB 7;  Length 669;
  Best Local Similarity   97.0%;  Pred. No. 1.1e-281;
  Matches  652;  Conservative    4;  Mismatches    7;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||| ||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297
```

```
Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669
```

```
RESULT 14
ADF16148
ID   ADF16148 standard; protein; 669 AA.
XX
AC   ADF16148;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1235.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1235; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 669 AA;

  Query Match             96.2%;  Score 3432.5;  DB 7;  Length 669;
  Best Local Similarity   97.0%;  Pred. No. 1.1e-281;
  Matches  652;  Conservative    4;  Mismatches    7;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         |||||| |||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHSEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669
```

```
RESULT 15
ADF16145
ID   ADF16145 standard; protein; 669 AA.
XX
AC   ADF16145;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1232.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1232; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 669 AA;

  Query Match             96.2%;  Score 3432.5;  DB 7;  Length 669;
  Best Local Similarity   97.0%;  Pred. No. 1.1e-281;
  Matches  652;  Conservative    4;  Mismatches    7;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :          |||||| |||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669


Search completed: April 19, 2006, 12:02:31
Job time : 162.884 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:02:52 ; Search time 30.6364 Seconds
(without alignments)
2116.769 Million cell updates/sec

Title: US-10-775-180-447
Perfect score: 3568
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....TCFAEEGKKLVAASQAALGL 674

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3108	87.1	609	1	ABHUS	serum albumin prec
2	2947	82.6	600	2	A47391	serum albumin prec
3	2627	73.6	608	2	S57632	serum albumin prec
4	2481.5	69.5	607	1	ABHOS	serum albumin prec
5	2451.5	68.7	607	1	ABBOS	serum albumin prec
6	2437.5	68.3	607	1	ABSHS	serum albumin prec
7	2431	68.1	608	1	ABRTS	serum albumin prec
8	2416.5	67.7	605	1	ABPGS	serum albumin prec
9	2387.5	66.9	609	2	JC5838	albumin - Mongolia
10	1861	52.2	453	2	A05139	serum albumin - mo
11	1562	43.8	615	1	ABCHS	serum albumin prec
12	1260.5	35.3	609	2	JC4258	alpha-fetoprotein
13	1256.5	35.2	609	1	FPHU	alpha-fetoprotein
14	1249.5	35.0	609	1	FPGO	alpha-fetoprotein
15	1207.5	33.8	607	1	ABXL72	74K albumin precur
16	1181.5	33.1	265	2	I46986	albumin - dog (fra
17	1175.5	32.9	608	1	ABXL68	68K serum albumin
18	1084	30.4	605	1	FPMS	alpha-fetoprotein
19	1067	29.9	611	1	FPRT	alpha-fetoprotein
20	1055	29.6	599	1	A54906	afamin precursor -
21	932.5	26.1	608	2	A53195	afamin precursor -
22	930	26.1	614	2	S59517	serum albumin prec
23	751.5	21.1	608	1	ABONS1	serum albumin 1 pr
24	746.5	20.9	608	1	ABONS2	serum albumin 2 pr
25	699	19.6	382	2	A37253	serum albumin - bu
26	440.5	12.3	1423	1	S27941	serum albumin - se
27	401	11.2	474	1	VYHUD	vitamin D-binding
28	400	11.2	476	1	VYRTD	vitamin D-binding
29	387	10.8	472	1	A35327	vitamin D-binding
30	247.5	6.9	180	1	GCBO	glucagon precursor
31	245.5	6.9	180	2	A57294	glucagon precursor
32	243.5	6.8	180	1	GCHU	glucagon precursor
33	243.5	6.8	180	1	GCRT	glucagon precursor
34	241.5	6.8	180	1	GCHY	glucagon precursor
35	240.5	6.7	180	1	GCGP	glucagon precursor
36	238.5	6.7	158	1	GCPG	glucagon precursor
37	230.5	6.5	180	1	GCRTDU	glucagon precursor
38	226.5	6.3	206	2	I51301	proglucagon - chic
39	225.5	6.3	151	1	GCCH	glucagon precursor
40	215.5	6.0	178	2	I51058	glucagon I precurs
41	209	5.9	101	1	GCFGB	glucagon precursor
42	207.5	5.8	178	2	I51057	glucagon II precur
43	200.5	5.6	122	1	GCAF2	glucagon 2 precurs
44	188.5	5.3	63	1	GCIDC	glucagon precursor
45	188	5.3	72	1	GCGXA	glucagon precursor

ALIGNMENTS

RESULT 1
ABHUS
serum albumin precursor [validated] - human
N;Alternate names: preproalbumin
N;Contains: kinetensin
C;Species: Homo sapiens (man)
C;Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C;Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S36
R;Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur
Nucleic Acids Res. 9, 6103-6114, 1981
A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli
A;Reference number: A93743; MUID:82081882; PMID:6171778
A;Accession: A93743
A;Molecule type: mRNA
A;Residues: 1-419,'K',421-609 <LAW>
A;Cross-references: UNIPROT:P02768; UNIPARC:UPI000002CE3A; EMBL:V00495; GB:J00078; GB:L0
R;Dugaiczyk, A.; Law, S.W.; Dennison, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A;Reference number: A93936; MUID:82105994; PMID:6275391
A;Accession: A93936
A;Molecule type: mRNA
A;Residues: 1-120,'G',122-609 <DUG>
A;Cross-references: UNIPARC:UPI0000156BE8; EMBL:V00494; NID:g28589; PIDN:CAA23753.1; PID
R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A;Reference number: I39427; MUID:86140099; PMID:2419329
A;Accession: I39427
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-26 <URA>
A;Cross-references: UNIPARC:UPI000002BD5F; GB:M13075; NID:g178330; PIDN:AAA51688.1; PID
R;Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family
A;Reference number: I59286; MUID:94181575; PMID:8134387
A;Accession: I59286
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 282-290,'KSRFDLQ' <WAT>
A;Cross-references: UNIPARC:UPI000011F7AF; GB:S69192; NID:g546032; PIDN:AAB30282.1; PID
A;Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R;Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl
A;Reference number: I59313; MUID:94294404; PMID:8022807
A;Accession: I59313
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 589-590,'ALPRRVKNLLLQVKLP' <MAD>
A;Cross-references: UNIPARC:UPI0000072EC4; GB:S70799; NID:g547231; PIDN:AAB31177.1; PID

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants
R;Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A;Reference number: G08292
A;Accession: G01747
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-120,'G',122-455 <MEN>
A;Cross-references: UNIPARC:UPI000016A1A8; EMBL:U22961; NID:g763428; PIDN:AAA64922.1; PI
R;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
A;Reference number: S55314; MUID:95275251; PMID:7755581
A;Accession: S55314
A;Molecule type: protein
A;Residues: 19-27 <LED>
A;Cross-references: UNIPARC:UPI00001743FA
R;Meloun, B.; Moravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A;Title: Complete amino acid sequence of human serum albumin.
A;Reference number: A91420; MUID:76187907; PMID:1225573
A;Accession: A91420
A;Molecule type: protein
A;Residues: 25-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',395-
A;Cross-references: UNIPARC:UPI00001743FB
R;Roehr, U.; Spiteller, G.; Tripier, D.
Justus Liebigs Ann. Chem. 9, 881-884, 1988
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from ur
A;Reference number: S06422
A;Note: this paper is in German, with an English abstract
A;Accession: S06422
A;Molecule type: protein
A;Residues: 25-48 <ROE>
A;Cross-references: UNIPARC:UPI0000052CDA
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A;Title: Mass spectrometric identification of modifications to human serum albumin treat
A;Reference number: S36882; MUID:93384321; PMID:8373198
A;Accession: S36882
A;Molecule type: protein
A;Residues: 45-67;141-160;311-337;469-490;570-581 <FIN>
A;Cross-references: UNIPARC:UPI00000423AC; UNIPARC:UPI00001743FC; UNIPARC:UPI00001743FD;
R;Kausler, E.; Spiteller, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A;Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol
A;Reference number: S17599; MUID:92126241; PMID:1772598
A;Accession: S17599
A;Molecule type: protein
A;Residues: 25-54;354-357;431-447 <KAU>
A;Cross-references: UNIPARC:UPI0000174400; UNIPARC:UPI0000174401; UNIPARC:UPI0000174402
A;Note: 49-Leu was also found
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: A45800
A;Molecule type: protein
A;Residues: 166-173 <CAR>
A;Cross-references: UNIPARC:UPI000004A560
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre
A;Reference number: A03239; MUID:86242180; PMID:3087352
A;Accession: A03239
A;Molecule type: protein
A;Residues: 166-173,'L' <MOG>
A;Cross-references: UNIPARC:UPI00000351D2
R;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A;Title: Mutations in genetic variants of human serum albumin found in Italy.
A;Reference number: A38255; MUID:91062352; PMID:2247440
A;Accession: C38255
A;Molecule type: protein
A;Residues: 76-111 <GAL1>
A;Cross-references: UNIPARC:UPI0000174403
A;Accession: B38255
A;Molecule type: protein
A;Residues: 82-105,'K',107-110 <GAL2>
A;Cross-references: UNIPARC:UPI0000174403
A;Note: this variant is designated albumin Vibo Valentia
A;Accession: A38255
A;Molecule type: protein
A;Residues: 76-83,'K',85-106 <GAL3>
A;Cross-references: UNIPARC:UPI0000174405
A;Note: this variant is designated albumin Torino
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin Her
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: protein
A;Residues: 255-263,'E',265-281 <MIN1>
A;Cross-references: UNIPARC:UPI0000174406
A;Note: this variant is designated albumin Herborn
R;Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,
Biochim. Biophys. Acta 1119, 232-238, 1992
A;Title: Two alloalbumins with identical electrophoretic mobility are produced by differ
A;Reference number: S21078; MUID:92190239; PMID:1347703
A;Accession: S21078
A;Molecule type: protein
A;Residues: 354-356,'K',358-378 <MIN2>
A;Cross-references: UNIPARC:UPI0000174407
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,
R;He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A;Title: Atomic structure and chemistry of human serum albumin.
A;Reference number: A46756; MUID:92334427; PMID:1630489
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,
A;Reference number: A94442
A;Contents: annotation; three-dimensional structure and disulfide bonds
R;Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A;Title: Disulfide bonds in human serum albumin.
A;Reference number: A90930
A;Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A;Reference number: A90299; MUID:78186630; PMID:656055
A;Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjoholm, I., eds., 11-20,
A;Title: Serum albumin: conformation and active sites.
A;Reference number: A94408
A;Contents: annotation; binding sites
R;Harper, M.E.; Dugaiczyk, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes
A;Reference number: A90028; MUID:83279982; PMID:6192711
A;Contents: annotation; gene position
R;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicyclic acid.
A;Reference number: A46755; MUID:76257808; PMID:955075
A;Contents: annotation
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicyclic acid
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phospha
A;Reference number: A56294; MUID:92183881; PMID:1544460
A;Contents: annotation
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p
atase activity
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized

lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
C;Comment: A large number of variants of human serum albumin have been described.
C;Genetics:
A;Gene: GDB:ALB
A;Cross-references: GDB:118990; OMIM:103600
A;Map position: 4q11-4q13

Query Match 87.1%; Score 3108; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 5.1e-197;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy         149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy         209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy         269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy         329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy         389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy         449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy         509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy         569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy         629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
               |||||||||||||||||||||||||||||||||||||||||||||
Db         564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

RESULT 2
A47391
serum albumin precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A47391
R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A;Title: cDNA and protein sequence of polymorphic macaque albumins that differ in biliru
A;Reference number: A47391; MUID:93211971; PMID:8460152
A;Contents: B/B homozygote
A;Accession: A47391
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-600 <WAT>
A;Cross-references: UNIPROT:Q28522; UNIPARC:UPI000001257C4; GB:M90463; NID:g342294; PIDN:
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)
C;Superfamily: serum albumin; serum albumin repeat homology
F;21-194/Domain: serum albumin repeat homology <SA1>
F;213-386/Domain: serum albumin repeat homology <SA2>
F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 82.6%; Score 2947; DB 2; Length 600;
Best Local Similarity 93.5%; Pred. No. 2e-186;
Matches 546; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

```
Qy          89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
               || |||||||||||||:|| |||:||:|||||||||:||||||||||||||||||||||
Db          16 RDTHKSEVAHRFKDLGEEHFKGLVLVAFSQYLQQCPFEEHVKLVNEVTEFAKTCVADESA 75

Qy         149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
               |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          76 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPPLVRPE 135

Qy         209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
               |||||||||||| |||||||||:|||||||||||||||| |||||| ||||||||||||||
Db         136 VDVMCTAFHDNEATFLKKYLYEVARRHPYFYAPELLFFAARYKAAFAECCQAADKAACLL 195

Qy         269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
               |||||||||||||||||||||||||:||||||||||||:||||||||||||||||||||
Db         196 PKLDELRDEGKASSAKQRLKCASLQKFGDRAFKAWAVARLSQKFPKAEFAEVSKLVTDLT 255

Qy         329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
               |||||||||||||||||||||||:||||||||||||||:|||||||||:|||||||||
Db         256 KVHTECCHGDLLECADDRADLAKYMCENQDSISSKLKECCDKPLLEKSHCLAEVENDEMP 315

Qy         389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
               |||||||||:|||||||||||||||||||||||||||||||||||||:||||||| || ||||
Db         316 ADLPSLAADYVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVMLLLRLAKAYEATLEK 375

Qy         449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
               |||||||||||||||||||:|||||||||:|||||||||||||||||||||||||||||
Db         376 CCAAADPHECYAKVFDEFQPLVEEPQNLVKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 435

Qy         509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
               ||||||||||||||:|||| |||||||||||||||:|||||||||||::|||||||
Db         436 TPTLVEVSRNLGKVGAKCCKLPEAKRMPCAEDYLSVVLNRLCVLHEKTPVSEKVTKCCTE 495

Qy         569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
               |||||||||||||:|| |||| |||||||||||:|||||||:|:|||||||||||||||
Db         496 SLVNRRPCFSALELDEAYVPKAFNAETFTFHADMCTLSEKEKQVKKQTALVELVKHKPKA 555

Qy         629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 672
               |||||| |||:|||||||||||||| |||||| | ||||||||
Db         556 TKEQLKGVMDNFAAFVEKCCKADDKEACFAEEGPKFVAASQAAL 599
```

RESULT 3
S57632
serum albumin precursor - cat
C;Species: Felis silvestris catus (domestic cat)
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: JC4660; S57632
R;Hilger, C.; Grigioni, F.; Hentges, F.
Gene 169, 295-296, 1996
A;Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A;Reference number: JC4660; MUID:96194824; PMID:8647469
A;Accession: JC4660
A;Molecule type: mRNA
A;Residues: 1-608 <HI2>
A;Cross-references: UNIPROT:P49064; UNIPARC:UPI00001257C2; EMBL:X84842; NID:g886484; PI
A;Experimental source: liver
C;Comment: This protein is the major protein component in plasma. It functions as a mul
ein has 35 conserved cysteine residues.
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: liver; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRP>
F;25-608/Product: serum albumin #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>

```
  Query Match             73.6%;  Score 2627;  DB 2;  Length 608;
  Best Local Similarity   80.1%;  Pred. No. 2.4e-165;
  Matches  483;  Conservative   53;  Mismatches   57;  Indels   10;  Gaps    1;

Qy         70 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 129
              |:|  |           |  |:||:||:|||| |||||:|: |||:||:|||||||||||
Db         15 SAYSRG----------VTRREAHQSEIAHRFNDLGEEHFRGLVLVAFSQYLQQCPFEDHV 64

Qy        130 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 189
              |||||||||| ||||:|| ||:|||| | ||||||||:||: |||||||| |:||||||
Db         65 KLVNEVTEFAKGCVADQSAANCEKSLHELLGDKLCTVASLRDKYGEMADCCEKKEPERNE 124

Qy        190 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 249
              ||||||||||    :|| || | ||||||:||: || ||||||||||||||||||::|:
Db        125 CFLQHKDDNPGFGQLVTPEADAMCTAFHENEQRFLGKYLYEIARRHPYFYAPELLYYAEE 184

Qy        250 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 309
              ||  |||||:|||||||| ||:| ||::  |||||:|||||||||||||||||:|||||
Db        185 YKGVFTECCEAADKAACLTPKVDALREKVLASSAKERLKCASLQKFGERAFKAWSVARLS 244

Qy        310 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 369
              |:|||||||:|||||| |:| ||||||||||||||||||||||||||||:||||||
Db        245 QKFPKAEFAEISKLVTDLAKIHKECCHGDLLECADDRADLAKYICENQDSISTKLKECCG 304

Qy        370 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 429
              ||:|||||||:||| ||:||||| || |||| |:|||| |||||||| |||||:||||:
Db        305 KPVLEKSHCISEVERDELPADLPPLAVDFVEDKEVCKNYQEAKDVFLGTFLYEYSRRHPE 364

Qy        430 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 489
              ||| ||||||| || |||||||  ||  ||| |||||||||||| ||:| ||||||:|||
Db        365 YSVSLLLRLAKEYEATLEKCCATDDPPACYAHVFDEFKPLVEEPHNLVKTNCELFEKLGE 424

Qy        490 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 549
              | |||||||||||||||||||||||:|||||||| ||||:|: ||||||||||:|
Db        425 YGFQNALLVRYTKKVPQVSTPTLVEVSRSLGKVGSKCCTHPEAERLSCAEDYLSVVLNRL 484

Qy        550 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 609
              ||||||||||:|||||||||||||||||:||||||||||:|||||||||:||| | |
Db        485 CVLHEKTPVSERVTKCCTESLVNRRPCFSALQVDETYVPKEFSAETFTFHADLCTLPEAE 544

Qy        610 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 669
              :|||||:|||||:|||||||:|||| || || :||:||| |:||| |||||| ||||||:|
Db        545 KQIKKQSALVELLKHKPKATEEQLKTVMGDFGSFVDKCCAAEDKEACFAEEGPKLVAAAQ 604

Qy        670 AAL 672
              |||
Db        605 AAL 607


RESULT 4
ABHOS
serum albumin precursor - horse
C;Species: Equus caballus (domestic horse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S34053
R;Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A;Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A;Reference number: S34053; MUID:93345495; PMID:8344282
A;Accession: S34053
A;Molecule type: mRNA
A;Residues: 1-607 <HOA>
A;Cross-references: UNIPROT:P35747; UNIPARC:UPI00001257C3; GB:X74045; NID:g399671; PIDN:
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F;263/Binding site: bilirubin (Lys) #status predicted

  Query Match             69.5%;  Score 2481.5;  DB 1;  Length 607;
  Best Local Similarity   74.6%;  Pred. No. 8.9e-156;
  Matches  450;  Conservative   70;  Mismatches   72;  Indels   11;  Gaps    2;

Qy         70 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 129
              |:|  |           |  || ||||:|||| ||||::|| |||:||:||||||||||
Db         15 SAYSRG----------VLRRDTHKSEIAHRFNDLGEKHFKGLVLVAFSQYLQQCPFEDHV 64

Qy        130 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 189
              |||||||||| | |||||||||||||||||||||||| ||||:|||| ||||||||
Db         65 KLVNEVTEFAKKCAADESAENCDKSLHTLFGDKLCTVATLRATYGELADCCEKQEPERNE 124

Qy        190 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 249
              ||| ||||:||||:| :|| |  | || :: : || |||||:|||||||| ||||| |:
Db        125 CFLTHKDDHPNLPKL-KPEPDAQCAAFQEDPDKFLGKYLYEVARRHPYFYGPELLFHAEE 183

Qy        250 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 309
              ||| ||||| | || |||:|||| |::    ||||:||||:| | ||||| |||:|||||
Db        184 YKADFTECCPADDKLACLIPKLDALKERILLSSAKERLKCSSFQNFGERAVKAWSVARLS 243

Qy        310 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 369
              |:||||:||||||:|||||| |||||||||||||||||||||||:||||| ||| ||:
Db        244 QKFPKADFAEVSKIVTDLTKVHKECCHGDLLECADDRADLAKYICEHQDSISGKLKACCD 303

Qy        370 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 429
              ||||:||||||||: |::|:|||:|||| | |::||:| :|||||| |||||:|||||
Db        304 KPLLQKSHCIAEVKEDDLPSDLPALAADFAEDKEICKHYKDAKDVFLGTFLYEYSRRHPD 363

Qy        430 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 489
              ||| ||||:||||| ||||||| |||   ||  |||:| |||||::|:|:|:||:|||:||
Db        364 YSVSLLLRIAKTYEATLEKCCAEADPPACYRTVFDQFTPLVEEPKSLVKKNCDLFEEVGE 423

Qy        490 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 549
              | |||||:|||||| ||||||||||: | ||||||:||| ||::|:||:|::|:: ||:|
Db        424 YDFQNALIVRYTKKAPQVSTPTLVEIGRTLGKVGSRCCKLPESERLPCSENHLALALNRL 483

Qy        550 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 609
              ||||||||||:::|||||:||  |||||||||:|| |||||| |||||||||||||| | |
Db        484 CVLHEKTPVSEKITKCCTDSLAERRPCFSALELDEGYVPKEFKAETFTFHADICTLPEDE 543

Qy        610 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 669
              :|||||:|| |||||||||||||| |: :|:||| |||  :||| |||||| ||||:||
Db        544 KQIKKQSALAELVKHKPKATKEQLKTVLGNFSAFVAKCCGREDKEACFAEEGPKLVASSQ 603

Qy        670 AAL 672
              ||
Db        604 LAL 606


RESULT 5
ABBOS
serum albumin precursor [validated] - bovine
N;Alternate names: 67K protein; preproalbumin
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94
R;Holowachuk, E.W.; Stoltenborg, J.K.; Reed, R.G.; Peters Jr., T.
submitted to the EMBL Data Library, August 1991
A;Description: Bovine serum albumin: cDNA sequence and expression.
A;Reference number: A38885
A;Accession: A38885
A;Molecule type: mRNA
A;Residues: 1-607 <HOL>
A;Cross-references: UNIPROT:P04277; UNIPARC:UPI0000174408; EMBL:M73215
R;Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
A;Title: Rapid confirmation and revision of the primary structure of bovine serum albumi
A;Reference number: A36401; MUID:91083649; PMID:2260975
A;Accession: A36401
```

A;Molecule type: protein
A;Residues: 25-41,'H',43-189,'E',191-213,'T',215-323,'D',325-393,'TS',396-607 <HIR>
A;Cross-references: UNIPARC:UPI0000174409
R;MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-485, 1979
A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A;Reference number: A91258; MUID:80024278; PMID:488109
A;Accession: A91258
A;Molecule type: protein
A;Residues: 1-32 <MAG>
A;Cross-references: UNIPARC:UPI000017440A
R;Hsieh, J.C.; Lin, F.P.; Tam, M.F.
Anal. Biochem. 170, 1-8, 1988
A;Title: Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing g
A;Reference number: A60808; MUID:88267456; PMID:3389500
A;Accession: B60808
A;Molecule type: protein
A;Residues: 25-41 <HSI>
A;Cross-references: UNIPARC:UPI000017440B
R;Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albu
A;Reference number: S10780; MUID:90336641; PMID:2379503
A;Accession: S10780
A;Molecule type: protein
A;Residues: 25-41,'H',43-57,59-64 <STR>
A;Cross-references: UNIPARC:UPI000017440C
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: D45800
A;Molecule type: protein
A;Residues: 163-172 <CAR>
A;Cross-references: UNIPARC:UPI000017440D
R;Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.
J. Biol. Chem. 262, 5968-5973, 1987
A;Title: Structure of a biologically active neurotensin-related peptide obtained from pe
A;Reference number: A26693; MUID:87194805; PMID:2437111
A;Accession: A26693
A;Molecule type: protein
A;Residues: 165-172,'L' <CA2>
A;Cross-references: UNIPARC:UPI00000351D2
R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.
Biochem. J. 191, 867-868, 1980
A;Title: Sequence of residues 400-403 of bovine serum albumin.
A;Reference number: A90309; MUID:82023364; PMID:7283978
A;Accession: A90309
A;Molecule type: protein
A;Residues: 402-433 <REE>
A;Cross-references: UNIPARC:UPI000017440E
R;Brown, J.R.
Fed. Proc. 34, 591, 1975
A;Title: Structure of bovine serum albumin.
A;Reference number: A91458
A;Accession: A91458
A;Molecule type: protein
A;Residues: 25-41,'H',43-117,'EQ',120-179,181-189,'E',191-194,'A',196-213,'T',215-288,'E
A;Cross-references: UNIPARC:UPI000017440F; UNIPARC:UPI0000174410
R;Brown, J.R.
submitted to the Atlas, April 1975
A;Reference number: A94551
A;Accession: A94551
A;Molecule type: protein
A;Residues: 190-195 <BR2>
A;Cross-references: UNIPARC:UPI0000174411
R;Brown, J.R.
Fed. Proc. 33, 1389, 1974
A;Reference number: A91457
A;Contents: annotation; disulfide bonds
R;Werlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC

A;Reference number: S55232; MUID:95031935; PMID:7945219
A;Accession: S55232
A;Status: preliminary
A;Molecule type: protein
A;Residues: 529-536;569-572 <WER>
A;Cross-references: UNIPARC:UPI0000174412; UNIPARC:UPI0000174413
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; copper binding; duplication; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-607/Product: serum albumin #status experimental <MPT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4

Query Match 68.7%; Score 2451.5; DB 1; Length 607;
Best Local Similarity 75.7%; Pred. No. 8.4e-154;
Matches 442; Conservative 71; Mismatches 70; Indels 1; Gaps 1;

```
Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              || ||||:|||||||||| || ||||||:||||||::|||||||:||||||||||||
Db         24 RDTHKSEIAHRFKDLGEEQFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFAKTCVADESH 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              |:|||||||||:|| ||:||||||:||||| |||||||||||| ||||:|:||:| :|:
Db         84 AGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKL-KPD 142

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
               : :|  |  :|: |  ||||||||||||||||||::| :|   | :|||| || ||||
Db        143 PNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQDCCQAEDKGACLL 202

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||:: :|::  ||||:|||:|||:|||||||| |||:||||||:|||||| ||:|||||||
Db        203 PKIETMREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKAEFVEVTKLVTDLT 262

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||| ||||||||||||||||||||:|||:|||||||||:|||||||||||||||| | :|
Db        263 KVHKECCHGDLLECADDRADLAKYICDNQDTISSKLKECCDKPLLEKSHCIAEVEKDAIP 322

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
               :|| | ||| | ||||||| |||| ||| |||||:||||:|:| :|||||| || |||:
Db        323 ENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYEYSRRHPEYAVSVLLRLAKEYEATLEE 382

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              |||  ||| ||: ||||: | ||:||||||||||: ||:|||| |||||:||||:||||||
Db        383 CCAKDDPHACYSTVFDKLKHLVDEPQNLIKQNCDQFEKLGEYGFQNALIVRYTRKVPQVS 442

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              |||||||||:|||||::||  ||::|||| |||||::||:||||||||||::|||||||
Db        443 TPTLVEVSRSLGKVGTRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTE 502

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              |||||||||||  ||||||| |: : |||||||||| : |:||||||||||:||||||
Db        503 SLVNRRPCFSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQTALVELLKHKPKA 562

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 672
              |:|||| ||::| |||:||| ||||| ||| || ||| ::| ||
Db        563 TEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLVVSTQTAL 606
```

RESULT 6
ABSHS
serum albumin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S06936
R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A;Reference number: S06936; MUID:90098888; PMID:2602160
A;Accession: S06936

```
A;Molecule type: mRNA
A;Residues: 1-607 <BRO>
A;Cross-references: UNIPROT:P14639; UNIPARC:UPI00001257CB; EMBL:X17055; NID:g1386; PIDN:
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F;263/Binding site: bilirubin (Lys) #status predicted

  Query Match             68.3%;  Score 2437.5;  DB 1;  Length 607;
  Best Local Similarity   75.0%;  Pred. No. 7e-153;
  Matches  438;  Conservative   73;  Mismatches   72;  Indels    1;  Gaps    1;

Qy          89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
               || ||||:|||| ||||||| |||||:||||||||::||||| |:||||||||||||
Db          24 RDTHKSEIAHRFNDLGEENFQGLVLIAFSQYLQQCPFDEHVKLVKELTEFAKTCVADESH 83

Qy         149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
                 |||||||||||:|| |||||||||:||||| |||||||||||| ||||:|:||:| :||
Db          84 AGCDKSLHTLFGDELCKVATLRETYGDMADCCEKQEPERNECFLNHKDDSPDLPKL-KPE 142

Qy         209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
               | :|  |  :|: |  |||||:||||||||||||::| :|   | ||||| || ||||
Db         143 PDTLCAEFKADEKKFWGKYLYEVARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLL 202

Qy         269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
               ||:| :|::  ||||:|||:|||:||||||| |||:||||||:||||:| :|:|:|||||
Db         203 PKIDAMREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKADFTDVTKIVTDLT 262

Qy         329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
               ||| ||||||||||||||||||||::||::|||||||:||:|||||||||: | :|
Db         263 KVHKECCHGDLLECADDRADLAKYICDHQDALSSKLKECCDKPVLEKSHCIAEVDKDAVP 322

Qy         389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
                :|| | ||| | |:|||| |||||||| |||||:||||:|:| :|||||| || |||
Db         323 ENLPPLTADFAEDKEVCKNYQEAKDVFLGSFLYEYSRRHPEYAVSVLLRLAKEYEATLED 382

Qy         449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
               |||  ||| ||| ||| | ||:|||||||:||||||: ||| |||||:||||:| ||||
Db         383 CCAKEDPHACYATVFDKLKHLVDEPQNLIKKNCELFEKHGEYGFQNALIVRYTRKAPQVS 442

Qy         509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
               ||||||:||:|||||:|||  ||::|||| |||||::||:|||||||||||::|||||||
Db         443 TPTLVEISRSLGKVGTKCCAKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTE 502

Qy         569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
               |||||||||| | :|||||| |: : |||||||||| : |:||||||||||:||||||
Db         503 SLVNRRPCFSDLTLDETYVPKPFDEKFFTFHADICTLPDTEKQIKKQTALVELLKHKPKA 562

Qy         629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 672
               | |||| ||::| |||:||| ||||| ||  || ||||::||||
Db         563 TDEQLKTVMENFVAFVDKCCAADDKEGCFVLEGPKLVASTQAAL 606


RESULT 7
ABRTS
serum albumin precursor - rat
N;Alternate names: preproalbumin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 09-Jul-2004
C;Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R;Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A;Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A;Reference number: A93872; MUID:81223722; PMID:7017712
A;Accession: A93872
A;Molecule type: mRNA
A;Residues: 1-608 <SAR>
A;Cross-references: UNIPROT:P02770; UNIPARC:UPI00001257CA; GB:V01222; GB:J00698; NID:g55
R;Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A;Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analys
A;Reference number: A92211; MUID:77249657; PMID:893447
A;Note: cleavages during protein maturation
A;Accession: A92211
A;Molecule type: protein
A;Residues: 1-38 <STR>
A;Cross-references: UNIPARC:UPI0000174416
R;Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A;Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavag
A;Reference number: A91946; MUID:78109429; PMID:564345
A;Accession: A91946
A;Molecule type: protein
A;Residues: 25-222 <IS1>
A;Cross-references: UNIPARC:UPI0000174417
R;Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A;Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino a
A;Reference number: A91940; MUID:76260153; PMID:956149
A;Accession: A91940
A;Molecule type: protein
A;Residues: 223-288;572-608 <IS2>
A;Cross-references: UNIPARC:UPI0000174418; UNIPARC:UPI0000174419
A;Note: 262-Leu was also found
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.
A;Reference number: A90758; MUID:79001617; PMID:80265
A;Contents: annotation; copper binding
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: C45800
A;Status: preliminary
A;Molecule type: protein
A;Residues: 166-173 <CAR>
A;Cross-references: UNIPARC:UPI0000017441A
R;Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A;Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved
A;Reference number: I57621; MUID:87286876; PMID:3475566
A;Accession: I57621
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: UNIPARC:UPI000011E8B8; GB:M16825; NID:g202828; PIDN:AAA40712.1; PID:
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-608/Product: serum albumin #status experimental <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status experimental
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4

  Query Match             68.1%;  Score 2431;  DB 1;  Length 608;
  Best Local Similarity   73.5%;  Pred. No. 1.9e-152;
  Matches  429;  Conservative   82;  Mismatches   73;  Indels    0;  Gaps    0;

Qy          89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
               |:|||||:|||||||||::|| |||||:||||:||:|:|:||| |||:|||||||||:|
Db          24 REAHKSEIAHRFKDLGEQHFKGLVLIAFSQYLQKCPYEEHIKLVQEVTDFAKTCVADENA 83

Qy         149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
```

```
Db         84 ENCDKSIHTLFGDKLCAIPKLRDNYGELADCCAKQEPERNECFLQHKDDNPNLPPFQRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268

Db        144 AEAMCTSFQENPTSFLGHYLHEVARRHPYFYAPELLYYAEKYNEVLTQCCTESDKAACLT 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328

Db        204 PKLDAVKEKALVAAVRQRMKCSSMQRFGERAFKAWAVARMSQRFPNAEFAEITKLATDVT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388

Db        264 KINKECCHGDLLECADDRAELAKYMCENQATISSKLQACCDKPVLQKSQCLAETEHDNIP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448

Db        324 ADLPSIAADFVEDKEVCKNYAEAKDVFLGTFLYEYSRRHPDYSVSLLLRLAKKYEATLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508

Db        384 CCAEGDPPACYGTVLAEFQPLVEEPKNLVKTNCELYEKLGEYGFQNAVLVRYTQKAPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568

Db        444 TPTLVEAARNLGRVGTKCCTLPEAQRLPCVEDYLSAILNRLCVLHEKTPVSEKVTKCCSG 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628

Db        504 SLVERRPCFSALTVDETYVPKEFKAETFTFHSDICTLPDKEKQIKKQTALAELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 672

Db        564 TEDQLKTVMGDFAQFVDKCCKAADKDNCFATEGPNLVARSKEAL 607
```

RESULT 8
ABPGS
serum albumin precursor - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S01382; A61006
R;Weinstock, J.; Baldwin, G.S.
Nucleic Acids Res. 16, 9045, 1988
A;Title: Nucleotide sequence of porcine liver albumin.
A;Reference number: S01382; MUID:89016582; PMID:3174440
A;Accession: S01382
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-605 <WEI>
A;Cross-references: UNIPROT:P08835; UNIPARC:UPI000001257C7; EMBL:X12422; NID:g1875; PIDN:
R;Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
J. Bone Miner. Res. 4, 235-241, 1989
A;Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral
A;Reference number: A61006; MUID:89269769; PMID:2728927
A;Accession: A61006
A;Molecule type: protein
A;Residues: 23-51,'X',53-54;'XXXGY',146,'E',148,'E',150-151,'XVN',155 <LIM>
A;Cross-references: UNIPARC:UPI0000174414; UNIPARC:UPI0000174415
A;Experimental source: dental enamel
A;Note: albumin and other serum proteins are also found in bone
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-16/Domain: signal sequence (fragment) #status predicted <SIG>
F;17-22/Domain: propeptide #status predicted <PRO>
F;23-605/Product: serum albumin #status predicted <MAT>
F;27-199/Domain: serum albumin repeat homology <SA1>
F;218-391/Domain: serum albumin repeat homology <SA2>
F;410-589/Domain: serum albumin repeat homology <SA3>
F;75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4
F;261/Binding site: bilirubin (Lys) #status predicted

Query Match 67.7%; Score 2416.5; DB 1; Length 605;
Best Local Similarity 76.1%; Pred. No. 1.7e-151;
Matches 439; Conservative 67; Mismatches 70; Indels 1; Gaps 1;

```
Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148

Db         22 RDTYKSEIAHRFKDLGEQYFKGLVLIAFSQHLQQCPYEEHVKLVREVTEFAKTCVADESA 81

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208

Db         82 ENCDKSIHTLFGDKLCAIPSLREHYGDLADCCEKEEPERNECFLQHKNDNPDIPKL-KPD 140

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268

Db        141 PVALCADFQEDEQKFWGKYLYEIARRHPYFYAPELLYYAIIYKDVFSECCQAADKAACLL 200

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328

Db        201 PKIEHLREKVLTSAAKQRLKCASIQKFGERAFKAWSLARLSQRFPKADFTEISKIVTDLA 260

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388

Db        261 KVHKECCHGDLLECADDRADLAKYICENQDTISTKLKECCDKPLLEKSHCIAEAKRDELP 320

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448

Db        321 ADLNPLEHDFVEDKEVCKNYKEAKDVFLGTFLYEYSRRHPDYSVSLLLRIAKIYEATLED 380

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508

Db        381 CCAKEDPPACYATVFDKFQPLVDEPKNLIKQNCELFEKLGEYGFQNALIVRYTKKVPQVS 440

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568

Db        441 TPTLVEVARKLGLVGSRCCKRPEEERLSCAEDYLSLVLNRLCVLHEKTPVSEKVTKCCTE 500

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628

Db        501 SLVNRRPCFSALTPDETYKPKEFVEGTFTFHADLCTLPEDEKQIKKQTALVELLKHKPHA 560

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLV 665

Db        561 TEEQLRTVLGNFAAFVQKCCAAPDHEACFAVEGPKFV 597
```

RESULT 9
JC5838
albumin - Mongolian jird
C;Species: Meriones unguiculatus (Mongolian jird)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5838
R;Yoshida, K.; Seto-Ohshima, A.; Sinohara, H.
DNA Res. 4, 351-354, 1997
A;Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in th
A;Reference number: JC5838; MUID:98116663; PMID:9455485
A;Accession: JC5838
A;Molecule type: mRNA
A;Residues: 1-609 <YOS>
A;Cross-references: UNIPROT:O35090; UNIPARC:UPI00001257C5; DDBJ:AB006197; NID:g2317277;
A;Experimental source: liver
C;Superfamily: serum albumin; serum albumin repeat homology
F;222-395/Domain: serum albumin repeat homology <SA2>

Query Match 66.9%; Score 2387.5; DB 2; Length 609;
Best Local Similarity 73.8%; Pred. No. 1.4e-149;
Matches 432; Conservative 65; Mismatches 87; Indels 1; Gaps 1;

```
Qy         89 RD-AHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADES 147

Db         24 RDAAHKSEIAHRYKDLGEKYFKGLVLYTFSQYLQKCSYEEHVKLVREVTDFASNCAKDES 83

Qy        148 AENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRP 207
```

```
                  ||||||||||||||||||::    | | ||||||||||||||||||||||||||||| ||    |
Db          84 AENCDKSLHTLFGDKLCSLPNFGEKYAEMADCCAKQEPERNECFLQHKDDNPQLPPFKRA 143

Qy         208 EVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACL 267
               | | ||||| :| | |:  ||:|:|||||||| ||||: | :| |  |||| | || |||
Db         144 EPDAMCTAFQENAEAFMGHYLHEVARRHPYFYGPELLYLADKYTAVLTECCAADDKGACL 203

Qy         268 LPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDL 327
                ||||  |:::   |: :|||||:|::|||||||||||||:|| || |:|||::|| |||
Db         204 TPKLDALKEKALVSAVRQRLKCSSMKKFGERAFKAWAVARMSQTFPNADFAEITKLATDL 263

Qy         328 TKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEM 387
               |||  |||||||||||||||:||||:|||| |||||: ||:| :|:|| |:||||:|:|
Db         264 TKVTQECCHGDLLECADDRAELAKYMCENQASISSKLQACCDKEMLQKSQCLAEVEHDDM 323

Qy         388 PADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLE 447
               |||||:| ||||| |||||||||||||||| |||||:||||:||| ||||||| || |||
Db         324 PADLPALTADFVEDKDVCKNYAEAKDVFLGTFLYEYSRRHPEYSVSLLLRLAKKYEATLE 383

Qy         448 KCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQV 507
               |||| |||| ||  |||||||||||||||:| ||||:|:|||| ||||:|||||| |||
Db         384 KCCAEADPHACYGHVFDEFKPLVEEPQNLVKSNCELYEKLGEYGFQNAVLVRYTKKAPQV 443

Qy         508 STPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCT 567
               ||||||| :|:||:||: ||  || ||:|| ||||| :||::|:|||||||::|||||:
Db         444 STPTLVEAARSLGRVGTHCCALPEKKRLPCVEDYLSAILNRVCLLHEKTPVSEQVTKCCS 503

Qy         568 ESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPK 627
               ||| |||||||| |||||||||| |||||||:|||| |||:|::|||| |||||||:
Db         504 GSLVERRPCFSALPVDETYVPKEFKAETFTFHANICTLPEKEKQMEKQTALAELVKHKPQ 563

Qy         628 ATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 672
               ||:|||| || ||| |:||||| :||| ||: || |||| || ||
Db         564 ATEEQLKKVMGDFAEFLEKCCKQEDKEACFSTEGPKLVAESQKAL 608


RESULT 10
A05139
serum albumin - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: A05139; I48638
R;Minghetti, P.P.; Law, S.W.; Dugaiczyk, A.
Mol. Biol. Evol. 2, 347-358, 1985
A;Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudog
A;Reference number: A93055; MUID:88216123; PMID:2452956
A;Accession: A05139
A;Molecule type: mRNA
A;Residues: 1-418 <MIN>
A;Cross-references: UNIPROT:P07724; UNIPARC:UPI000016CBE6; GB:M16111; NID:g191764; PIDN:
R;Boccaccio, C.; Deschatrette, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A;Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the
A;Reference number: I48638; MUID:90269606; PMID:1971802
A;Accession: I48638
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 379-453 <BOC>
A;Cross-references: UNIPARC:UPI000016CEAE; EMBL:X13060; NID:g52939; PIDN:CAA31458.1; PID
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F;123-296/Domain: serum albumin repeat homology <SA2>
F;315-453/Domain: serum albumin repeat homology (fragment) <SA3>

  Query Match             52.2%;  Score 1861;  DB 2;  Length 453;
  Best Local Similarity   72.2%;  Pred. No. 4.4e-115;
  Matches  327;  Conservative   64;  Mismatches   62;  Indels    0;  Gaps    0;

Qy         164 CTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETF 223
               | :  ||| |||:|||| |||||||||||||||||:||   ||| : |||:| :|  ||
Db           1 CAIPNLRENYGELADCCTKQEPERNECFLQHKDDNPSLPPFERPEAEAMCTSFKENPTTF 60

Qy         224 LKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSA 283
               :  ||:|:|||||||||||::|::|    |:||  ||| :|| |||| ::::   ||
Db          61 MGHYLHEVARRHPYFYAPELLYYAEQYNEILTQCCAEADKESCLTPKLDGVKEKALVSSV 120

Qy         284 KQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECA 343
               :||:||:|:|||||||||||||||| || |:|||::|| ||||||: ||||||||||
Db         121 RQRMKCSSMQKFGERAFKAWAVARLSQTFPNADFAEITKLATDLTKVNKECCHGDLLECA 180

Qy         344 DDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKD 403
               ||||:||||:|||| :|||||: ||:|||||:|:||::|||:| ||||||::|||||| ::
Db         181 DDRAELAKYMCENQATISSKLQTCCDKPLLKKAHCLSEVEHDTMPADLPAIAADFVEDQE 240

Qy         404 VCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVF 463
               |||||||||||||| ||||||||||||| ||||||| || ||||||| |:|  ||  |
Db         241 VCKNYAEAKDVFLGTFLYEYSRRHPDYSVSLLLRLAKKYEATLEKCCAEANPPACYGTVL 300

Qy         464 DEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVG 523
               ||:||||||:||:| ||:|:|:|||| ||||:|||||:| |||||||||| :||||:||
Db         301 AEFQPLVEEPKNLVKTNCDLYEKLGEYGFQNAILVRYTQKAPQVSTPTLVEAARNLGRVG 360

Qy         524 SKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVD 583
               :|||  || :|:|| ||||| :||::|:|||||||: |||||: ||| |||||||| ||
Db         361 TKCCTLPEDQRLPCVEDYLSAILNRVCLLHEKTPVSEHVTKCCSGSLVERRPCFSALTVD 420

Qy         584 ETYVPKEFNAETFTFHADICTLSEKERQIKKQT 616
               |||||||| |||||||:||||| |||:||||||
Db         421 ETYVPKEFKAETFTFHSDICTLPEKEKQIKKQT 453


RESULT 11
ABCHS
serum albumin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S15571; A05078; A13451
R;Cassady, A.I.; Salklld, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A;Reference number: S15571
A;Accession: S15571
A;Molecule type: mRNA
A;Residues: 1-615 <CAS>
A;Cross-references: UNIPROT:P19121; UNIPARC:UPI00001257C1; EMBL:X60688; NID:g63747; PIDN
R;Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A;Title: The 5' noncoding and flanking regions of the avian very low density apolipoprot
A;Reference number: A05078; MUID:83161037; PMID:6187737
A;Accession: A05078
A;Molecule type: DNA
A;Residues: 1-28 <HAC>
A;Cross-references: UNIPARC:UPI000017128E; GB:V00381; NID:g63038; PIDN:CAA23680.1; PID:g
R;Rosen, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A;Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A;Reference number: A13451; MUID:78019943; PMID:911327
A;Accession: A13451
A;Molecule type: protein
A;Residues: 19-23,'M',25-30 <ROS>
A;Cross-references: UNIPARC:UPI000017441B
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thyr
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-26/Domain: propeptide #status predicted <PRO>
F;27-613/Product: serum albumin #status predicted <MAT>
F;32-206/Domain: serum albumin repeat homology <SA1>
F;225-398/Domain: serum albumin repeat homology <SA2>
F;417-596/Domain: serum albumin repeat homology <SA3>
F;30/Binding site: copper (His) #status predicted
F;80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-397,

  Query Match             43.8%;  Score 1562;  DB 1;  Length 615;
  Best Local Similarity   46.9%;  Pred. No. 3.1e-95;
```

```
  Matches  276;  Conservative  118;  Mismatches  192;  Indels    2;  Gaps    2;

Qy         89 RDA-HKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADES 147
              ||| |||| :|||: || || ||| : :| |||||:| :|    ||| :| : |: |||:|
Db         26 RDAEHKSEIAHRYNDLKEETFKAVAMITFAQYLQRCSYEGLSKLVKDVVDLAQKCVANED 85

Qy        148 AENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPR-LVR 206
              |  | | | ::  |::| |  ||::|| |||||:| :|||||||  |    |: :   |
Db         86 APECSKPLPSIILDEICQVEKLRDSYGAMADCCSKADPERNECFLSFKVSQPDFVQPYQR 145

Qy        207 PEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC 266
              |  ||:|  : ||  :||  ::| :|||||: ||| :| ||  :: |    ||: :|  ||
Db        146 PASDVICQEYQDNRVSFLGHFIYSVARRHPFLYAPAILSFAVDFEHALQSCCKESDVGAC 205

Qy        267 LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTD 326
              |  |   :|:: |  | ||:  |  |::||:| |:|  :  |||::||| |:|||| | |
Db        206 LDTKEIVMREKAKGVSVKQQYFCGILKQFGDRVFQARQLIYLSQKYPKAPFSEVSKFVHD 265

Qy        327 LTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDE 386
                 || ||| ||:|| || | :  :|  ||  | |:|:|||||::|:| || | | ||
Db        266 SIGVHKECCEGDMVECMDDMARMMSNLCSQQDVFSGKIKDCCEKPIVERSQCIMEAEFDE 325

Qy        387 MPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTL 446
              |||||||   ::| |:|||::    | |:  |:|||:||||::|: |::|:|| ||: |
Db        326 KPADLPSLVEKYIEDKEVCKSFEAGHDAFMAEFVYEYSRRHPEFSIQLIMRIAKGYESLL 385

Qy        447 EKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQ 506
              ||||   :| ||||   ::    ::| |:::| ||:|    ||  |  ::|:|||||:||
Db        386 EKCCKTDNPAECYANAQEQLNQHIKETQDVVKTNCDLLHDHGEADFLKSILIRYTKKMPQ 445

Qy        507 VSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC 566
              | |  |:|  : :  :|:|||:  | :|| |:| |||:|::  |   | ||::| |::||
Db        446 VPTDLLLETGKKMTTIGTKCCQLGEDRRMACSEGYLSIVIHDTCRKQETTPINDNVSQCC 505

Qy        567 TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKP 626
              ::   ||||||:|: ||  |||  || : |:|   :|:   :||:: :   |: |:| ||
Db        506 SQLYANRRPCFTAMGVDTKYVPPPFNPDMFSFDEKLCSAPAEERVEGQMKLLINLIKRKP 565

Qy        627 KATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              : |:||:| : | | | |:|||  |  ||| |||  |:  |:| ||:
Db        566 QMTEEQIKTIADGFTAMVDKCCKQSDINTCFGEEGANLIVQSRATLGI 613


RESULT 12
JC4258
alpha-fetoprotein precursor - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
C;Accession: JC4258
R;Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczyk, A.
Gene 162, 213-220, 1995
A;Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to t
A;Reference number: JC4258; MUID:96032345; PMID:7557431
A;Accession: JC4258
A;Molecule type: DNA
A;Residues: 1-609 <NIS>
A;Cross-references: UNIPROT:Q28789; UNIPARC:UPI000012A6FA; GB:U21916; NID:g841311; PIDN:
C;Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
o similar properties and structure.
C;Genetics:
A;Gene: afp
A;Map position: 3p
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-609/Product: alpha-fetoprotein #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;42,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             35.3%;  Score 1260.5;  DB 2;  Length 609;

  Best Local Similarity   39.7%;  Pred. No. 2.1e-75;
  Matches  238;  Conservative  117;  Mismatches  238;  Indels    7;  Gaps    3;

Qy         81 FIAWLVKGRDAHKSE-----VAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEV 135
              |:    : |  |::|     :   ::   | |   |  | |||::|:  :::  |:| :
Db         11 FLLNFTESRTLHRNEYGIASILDSYQCTAEINLTDLATIFFAQFVQEATYKEVSKMVKDA 70

Qy        136 TEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHK 195
                :    || :  | ::    | ::||    : | ||  :|||:: |  |: ||| ||
Db         71 LTAIEKPTGDEQSAGCLENQLPAFLEELCREKEILEKYGH-SDCCSQSEEGRHNCFLAHK 129

Qy        196 DDNP-NLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAF 254
                | ::|    ||    | |: :: |||: |::|||||||: ||| :| :| ||
Db        130 KPTPASIPFFQVPEPVTSCEAYEEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKII 189

Qy        255 TECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPK 314
                ||:|  :   |   |   :  | :||   :  || :: || | |:|  | :|||:| |
Db        190 PSCCKAENAVECFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFQAITVTKLSQKFTK 249

Qy        315 AEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLE 374
                | |: ||| |:  ||  || ||:|:|  |   :  |||  ||::|:|: |||:   ||
Db        250 VNFTEIQKLVLDVAHVHEHCCRGDVLDCLQDGEKIMSYICSQQDTLSNKITECCKLTTLE 309

Qy        375 KSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVL 434
              :  ||   |||| |  |      |:  :|  :  :  |::||  |::||:|||  :| :
Db        310 RGQCIIHAENDEKPEGLSPNLNRFLGDRDFNQFSSGEKNIFLASFVHEYSRRHPQLAVSV 369

Qy        435 LLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQN 494
              :||:|| |:  ||||     :| ||  |  :| :  ::| | | |::| ||::||||  ||
Db        370 ILRVAKGYQELLEKCFQTENPLECQDKGEEELQKYIQESQALAKRSCGLFQKLGEYYLQN 429

Qy        495 ALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHE 554
              | || |||| ||:::  |: ::| :    : ||:  | | : | |    :::  ||: ||
Db        430 AFLVAYTKKAPQLTSSELMAITRKMAATAATCCQLSEDKLLACGEGAADIIIGHLCIRHE 489

Qy        555 KTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKK 614
              |||:  | :||| |  |||||||:| |||||||  |: : | || |:|       |  |
Db        490 TTPVNPGVGQCCTSSYANRRPCFSSLVVDETYVPPAFSDDKFIFHKDLCQAQGVALQTMK 549

Qy        615 QTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |  |: ||| ||: |:|||:||: ||:  :||||:  ::| |||||||:|:: ::||||:
Db        550 QEFLINLVKQKPQITEEQLEAVIADFSGLLEKCCQGQEQEVCFAEEGQKLISKTRAALGV 609


RESULT 13
FPHU
alpha-fetoprotein precursor [validated] - human
N;Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C;Species: Homo sapiens (man)
C;Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 09-Jul-2004
C;Accession: A26624; S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042; A0
R;Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczyk, A.
Biochemistry 26, 1332-1343, 1987
A;Title: Structure, polymorphism, and novel repeated DNA elements revealed by a complete
A;Reference number: A26624; MUID:87185438; PMID:2436661
A;Accession: A26624
A;Molecule type: DNA
A;Residues: 1-609 <GIB>
A;Cross-references: UNIPROT:P02771; UNIPARC:UPI00000012A9; GB:M16110; NID:g773678; PIDN
R;McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilghman, S.; Krumlauf
Hum. Mol. Genet. 2, 379-384, 1993
A;Title: A G->A substitution in an HNF I binding site in the human alpha-fetoprotein gen
A;Reference number: S37655; MUID:93278385; PMID:7684942
A;Accession: S37655
A;Molecule type: DNA
A;Residues: 1-28 <MCV>
A;Cross-references: UNIPARC:UPI000016A4DF; EMBL:Z19532; NID:g28527; PIDN:CAA79592.1; PI
A;Note: the authors translated the codon TAT for residue 26 as Thr
R;Morinaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A;Title: Primary structures of human alpha-fetoprotein and its mRNA.
A;Reference number: A93961; MUID:83273664; PMID:6192439
A;Accession: A93961
```

A;Molecule type: mRNA
A;Residues: 1-609 <MOR>
A;Cross-references: UNIPARC:UPI00000012A9; GB:J00077; NID:g311348; PIDN:CAA24758.1; PID:
R;Beattie, W.G.; Dugaiczyk, A.
Gene 20, 415-422, 1982
A;Title: Structure and evolution of human alpha-fetoprotein deduced from partial sequenc
A;Reference number: A91497; MUID:83158778; PMID:6187626
A;Accession: A91497
A;Molecule type: mRNA
A;Residues: 429-556 <BEA>
A;Cross-references: UNIPARC:UPI0000174421; GB:J00076
R;Pucci, P.; Siciliano, R.; Malorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Terrana
Biochemistry 30, 5061-5066, 1991
A;Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A;Reference number: A23699; MUID:91242409; PMID:1709810
A;Accession: A23699
A;Molecule type: protein
A;Residues: 19-45;60-97;102-107;122-184;187-249;255-489;507-609 <PUC>
A;Cross-references: UNIPARC:UPI0000174422; UNIPARC:UPI0000174423; UNIPARC:UPI0000174424;
R;Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A;Title: Characterization of in vitro expressed human alpha-fetoprotein as highly reprod
A;Reference number: A61480; MUID:91225826; PMID:1709209
A;Accession: A61480
A;Molecule type: protein
A;Residues: 19-45;63-97;102-107;122-184;187-249;255-489;507-609 <TEC>
A;Cross-references: UNIPARC:UPI0000174422; UNIPARC:UPI0000174424; UNIPARC:UPI0000174425;
R;Yachnin, S.; Hsu, R.; Heinrikson, R.L.; Miller, J.B.
Biochim. Biophys. Acta 493, 418-428, 1977
A;Title: Studies on human alpha-fetoprotein. Isolation and characterization of monomeric
A;Reference number: A90624; MUID:77242506; PMID:70228
A;Accession: A90624
A;Molecule type: protein
A;Residues: 'S',20-22,'S',24-35 <YAC>
A;Cross-references: UNIPARC:UPI000017442A
A;Note: dimeric and trimeric forms have been found in addition to the monomeric form
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 37, 3663-3667, 1977
A;Title: Comparative chemical structure of human alpha-fetoproteins from fetal serum and
A;Reference number: A90757; MUID:78001760; PMID:71198
A;Accession: A90757
A;Molecule type: protein
A;Residues: 'S',20-30,'A',32-37,'A' <AOY>
A;Cross-references: UNIPARC:UPI000017442B
R;Ruoslahti, E.; Pihko, H.; Vaheri, A.; Seppala, M.; Virolainen, M.; Konttinen, A.
Johns Hopkins Med. J. Suppl. 3, 249-255, 1974
A;Title: 20. Alpha fetoprotein: structure and expression in man and inbred mouse strains
A;Reference number: A93042; MUID:75018719; PMID:4138095
A;Accession: A93042
A;Molecule type: protein
A;Residues: 'S',20-24,'Q',26-30,'A',32-35,'E',37-39 <RUO>
A;Cross-references: UNIPARC:UPI000017442C
R;Sakai, M.; Morinaga, T.; Urano, Y.; Watanabe, K.; Wegmann, T.G.; Tamaoki, T.
J. Biol. Chem. 260, 5055-5060, 1985
A;Title: The human alpha-fetoprotein gene. Sequence organization and the 5' flanking reg
A;Reference number: A92520; MUID:85182629; PMID:2580830
A;Contents: annotation; gene, exons and introns
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.
A;Reference number: A90758; MUID:79001617; PMID:80265
A;Contents: annotation; metal binding
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 39, 3571-3574, 1979
A;Title: alpha-Fetoprotein as a carrier protein in plasma and its bilirubin-binding abil
A;Reference number: A90759; MUID:80001710; PMID:89900
A;Contents: annotation; bilirubin binding
C;Comment: AFP is synthesized by the fetal liver and yolk sac. It occurs in the plasma c
o trace amounts after birth. The serum level in adults is usually less than 40 ng/ml. AF
C;Comment: Human AFP binds copper, nickel, and fatty acids as well as, and the bilirubin
properties.
C;Genetics:
A;Gene: GDB:AFP
A;Cross-references: GDB:119660; OMIM:104150
A;Map position: 4q11-4q13
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-609/Product: alpha-fetoprotein #status experimental <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;22/Binding site: copper (His) #status experimental
F;99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
F;249/Binding site: bilirubin (Lys) #status predicted
F;251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.2%; Score 1256.5; DB 1; Length 609;
Best Local Similarity 39.5%; Pred. No. 3.9e-75;
Matches 237; Conservative 118; Mismatches 238; Indels 7; Gaps 3;

```
Qy         81 FIAWLVKGRDAHKSE-----VAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEV 135
              |:    : |  |::|     :   ::   | :   |  | |||::|:  :::  |:| :
Db         11 FLLNFTESRTLHRNEYGIASILDSYQCTAEISLADLATIFFAQFVQEATYKEVSKMVKDA 70

Qy        136 TEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHK 195
                :    || :  | ::    | ::||    : | ||  :|||:: |  |: ||| ||
Db         71 LTAIEKPTGDEQSSGCLENQLPAFLEELCHEKEILEKYGH-SDCCSQSEEGRHNCFLAHK 129

Qy        196 DDNP-NLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAF 254
              | ::|   ||   | |: :: |||: |::|||||||: ||| :| :| ||
Db        130 KPTPASIPLFQVPEPVTSCEAYEEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKII 189

Qy        255 TECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPK 314
              ||:| :   |   |   :  | : ||   :  || :: || | |:|  | :|||:| |
Db        190 PSCCKAENAVECFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFQAITVTKLSQKFTK 249

Qy        315 AEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLE 374
              | |: ||| |:  ||  || ||:|:|  |   :  |||  ||::|:|: |||:   ||
Db        250 VNFTEIQKLVLDVAHVHEHCCRGDVLDCLQDGEKIMSYICSQQDTLSNKITECCKLTTLE 309

Qy        375 KSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVL 434
              :  ||   |||| |  |      |:  :|  :  :  |::||  |::||:|||| :| :
Db        310 RGQCIIHAENDEKPEGLSPNLNRFLGDRDFNQFSSGEKNIFLASFVHEYSRRHPQLAVSV 369

Qy        435 LLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQN 494
              :||:|| |:  ||||    :| ||  |  :| :  ::| | | |::| ||::|||| ||
Db        370 ILRVAKGYQELLEKCFQTENPLECQDKGEEELQKYIQESQALAKRSCGLFQKLGEYYLQN 429

Qy        495 ALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHE 554
              | || |||| ||::  |: ::| :    : ||:  | | : | |    :::  ||: ||
Db        430 AFLVAYTKKAPQLTSSELMAITRKMAATAATCCQLSEDKLLACGEGAADIIIGHLCIRHE 489

Qy        555 KTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKK 614
              |||:  | :||| |  |||||||:| |||||||  |: : | || |:|       |  |
Db        490 MTPVNPGVGQCCTSSYANRRPCFSSLVVDETYVPPAFSDDKFIFHKDLCQAQGVALQTMK 549

Qy        615 QTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |  |: ||| ||: |:|||:||: ||:  :||||:  ::| ||||||:||:: ::||||:
Db        550 QEFLINLVKQKPQITEEQLEAVIADFSGLLEKCCQGQEQEVCFAEEGQKLISKTRAALGV 609
```

RESULT 14
FPGO
alpha-fetoprotein precursor - gorilla
C;Species: Gorilla gorilla (gorilla)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A37970
R;Ryan, S.C.; Zielinski, R.; Dugaiczyk, A.
Genomics 9, 60-72, 1991
A;Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primates.
A;Reference number: A37970; MUID:91169517; PMID:1706310
A;Accession: A37970
A;Molecule type: DNA
A;Residues: 1-609 <RYA>

```
A;Cross-references: UNIPROT:P28050; UNIPARC:UPI000012A6F8; GB:M38272; NID:g817963; PIDN:
C;Genetics:
A;Map position: 4q11-12
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-609/Product: alpha-fetoprotein #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;22/Binding site: copper (His) #status predicted
F;99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
F;249/Binding site: bilirubin (Lys) #status predicted
F;251/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             35.0%;  Score 1249.5;  DB 1;  Length 609;
  Best Local Similarity   39.2%;  Pred. No. 1.1e-74;
  Matches  235;  Conservative  119;  Mismatches  239;  Indels    7;  Gaps    3;

Qy          81 FIAWLVKGRDAHKSE-----VAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEV 135
               |:    : |  |::|     :   ::   | :   |  | |||::|:  :::  |:| :
Db          11 FLLNFTESRTLHRNEYGIASILDSYQCTAEISLADLATIFFAQFVQEATYKEVSKMVKDA 70

Qy         136 TEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHK 195
                 :   || :  | ::    | ::||    : | || ::|||::  |  |: ||| ||
Db          71 LTAIEKPTGDEQSAGCLENQLPAFLEELCHEKEILEKYG-LSDCCSQSEEGRHNCFLAHK 129

Qy         196 DDNP-NLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAF 254
                | ::|    ||    | |: :: |||: |::|||||||: ||| :| :| ||
Db         130 KPTPASIPLFQVPEPVTSCEAYEEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKII 189

Qy         255 TECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPK 314
               ||:| :   |   |   :  | : ||   :  || :: || | |:|  | :|||:| |
Db         190 PSCCKAENAVECFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFQAITVTKLSQKFTK 249

Qy         315 AEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLE 374
               | |: ||| |:  ||  || ||:|:|  |   :  |||  ||::|:|: |||:   ||
Db         250 VNFTEIQKLVLDVAHVHEHCCRGDVLDCLQDGEKIMSYICSQQDTLSNKITECCKLTTLE 309

Qy         375 KSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVL 434
               :  ||   |||| |  |      |: :|  :  :  |::||  |::||:|||| :| :
Db         310 RGQCIIHAENDEKPEGLSPNLNRFLGDRDFNQFSSGEKNIFLASFVHEYSRRHPQLAVSV 369

Qy         435 LLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQN 494
               :||:|| |:  ||||     :| ||  |  :| :  ::| | | |::| ||::|||  ||
Db         370 ILRVAKGYQELLEKCFQTENPLECQDKGEEELQKYIQESQALAKRSCGLFQKLGEYYLQN 429

Qy         495 ALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHE 554
               | || |||| ||::  |: ::| :    : ||:  | | : | |    :::  ||: ||
Db         430 AFLVAYTKKAPQLTSSELMAITRKMAATAATCCQLSEDKLLACGEGAADIIIGHLCIRHE 489

Qy         555 KTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKK 614
               |||:  | :||| |  |||||||:| |||||||  |: : | || |:|      |  |
Db         490 MTPVNPGVGQCCTSSYANRRPCFSSLVVDETYVPPAFSDDKFIFHKDLCQAQGVALQTMK 549

Qy         615 QTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
               |  |: ||| ||: |:|||: |: ||:  :||||:  ::| ||||||:||:: :: |||:
Db         550 QEFLINLVKQKPQITEEQLETVIADFSGLLEKCCQGQEQEVCFAEEGQKLISKTRTALGV 609


RESULT 15
ABXL72
74K albumin precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: B41682; S02693; A05288
R;Moskaitis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A;Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic ac
e during development.
A;Reference number: A41682; MUID:89313788; PMID:2747653
A;Accession: B41682
A;Molecule type: mRNA
A;Residues: 3-607 <MOS>
A;Cross-references: UNIPROT:P14872; UNIPARC:UPI00001714FB; GB:M21442; NID:g213930; PIDN:
R;Schorpp, M.; Doebbeling, U.; Wagner, U.; Ryffel, G.U.
J. Mol. Biol. 199, 83-93, 1988
A;Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Dele
A;Reference number: S02692; MUID:88172470; PMID:2451026
A;Accession: S02693
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-48 <SCH>
A;Cross-references: UNIPARC:UPI000017441E; EMBL:Z26826
R;Wolffe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J.
Eur. J. Biochem. 146, 489-496, 1985
A;Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilizatio
A;Reference number: A05288; MUID:85126974; PMID:3971963
A;Accession: A05288
A;Molecule type: mRNA
A;Residues: 459-502,'L',504-557 <WOL>
A;Cross-references: UNIPARC:UPI000017441F; GB:M28276
A;Note: the authors translated the codon TAT for residue 63 as Thr
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thyr
C;Genetics:
A;Introns: 27/1
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: 74K serum albumin #status predicted <MAT>
F;32-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;30/Binding site: copper (His) #status predicted
F;80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,
F;256/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             33.8%;  Score 1207.5;  DB 1;  Length 607;
  Best Local Similarity   38.5%;  Pred. No. 6.5e-72;
  Matches  232;  Conservative  113;  Mismatches  243;  Indels   15;  Gaps    4;

Qy          70 SSYLEGQAAKEFIAWLVKGRDA---HKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFE 126
               ||::| :        ::  ||    |  :|  :  | |  || | |   :| ||:|  |
Db          13 SSFIESR--------ILFKRDTDADHHKHIADVYTALTERTFKGLTLAIVSQNLQKCSLE 64

Qy         127 DHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPE 186
               :  |||||: :|||:|: |:: | |:| : ||| ||||    :   |    :|||||:||
Db          65 ELSKLVNEINDFAKSCINDKTPE-CEKPVGTLFFDKLCADPAVGVNYEWSKECCAKQDPE 123

Qy         187 RNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFF 246
               | :||  |:|         ::|| :  |    :: :  |  :::| || ||  | | :|
Db         124 RAQCFKAHRDHEHT---SIKPEPEETCKLLKEHPDDLLSAFIHEEARNHPDLYPPAVLAL 180

Qy         247 AKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVA 306
               |:|      ||:  ||  |  |: :|  : :   ||   |  | | |:  ||  :|
Db         181 TKQYHKLAEHCCEEEDKEKCFSEKMKQLMKQSHSIEDKQHHFCWILDNFPEKVLKALNLA 240

Qy         307 RLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKE 366
               |:| |:|||||         ::|    :||| |: ||  :| :| :: |:::| :||||::
Db         241 RVSHRYPKAEFKLAHNFTEEVTHFIKDCCHDDMFECMTERLELTEHTCQHKDELSSKLEK 300

Qy         367 CCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARR 426
               ||  ||||:::||  :|||::||:|      :| |    ||: |||  :|||| :|:  :|:
Db         301 CCNIPLLERTYCIVTLENDDVPAELSQPITEFTEDPHVCEKYAENNEVFLGRYLHAVSRK 360

Qy         427 HPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQ 486
               | : |   ||: || ||: | |||   :| |||    | |    :|    :||||::  :
Db         361 HQELSEQFLLQSAKEYESLLNKCCKTDNPPECYKDGADRFMNEAKERFAYLKQNCDILHE 420

Qy         487 LGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVL 546
               ||| |:| ||:|||||:|||| ||: ::  :  :|  ||  || :|||||  |::::
Db         421 HGEYLFENELLIRYTKKMPQVSDETLIGIAHQMADIGEHCCAVPENQRMPCAEGDLTILI 480
```

```
Qy        547 NQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLS 606
              ::|   :|| ::: |  |||:|    | ||:||  || |||     :|| |   ||| :
Db        481 GKMCERQKKTFINNHVAHCCTDSYSGMRSCFTALGPDEDYVPPPVTDDTFHFDDKICTAN 540

Qy        607 EKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVA 666
              :||:|  ||  ||:|:|  ||  |  :    :|   |:||| ||: : ||  |   |:
Db        541 DKEKQHIKQKFLVKLIKVSPKLEKNHIDECSAEFLKMVQKCCTADEHQPCFDTEKPVLIE 600

Qy        667 ASQ 669
                |
Db        601 HCQ 603
```

Search completed: April 19, 2006, 12:09:59
Job time : 32.6364 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 11:57:02 ; Search time 178.074 Seconds
(without alignments)
2670.387 Million cell updates/sec

Title: US-10-775-180-447
Perfect score: 3568
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....TCFAEEGKKLVAASQAALGL 674

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3108	87.1	609	1	ALBU_HUMAN	P02768 homo sapien
2	3108	87.1	609	2	Q645G4_HUMAN	Q645g4 homo sapien
3	3108	87.1	609	2	Q5NVH5_PONPY	Q5nvh5 pongo pygma
4	3084	86.4	609	2	Q68DN5_HUMAN	Q68dn5 homo sapien
5	3073	86.1	609	2	Q56G89_HUMAN	Q56g89 homo sapien
6	3066	85.9	627	2	Q5D0D7_HUMAN	Q5d0d7 homo sapien
7	2947	82.6	600	1	ALBU_MACMU	Q28522 macaca mula
8	2627	73.6	608	1	ALBU_FELCA	P49064 felis silve
9	2574	72.1	608	1	ALBU_CANFA	P49822 canis famil
10	2509	70.3	608	2	Q95VB7_SCHMA	Q95vb7 schistosoma
11	2501.5	70.1	607	1	ALBU_EQUAS	Q5xle4 equus asinu
12	2481.5	69.5	607	1	ALBU_HORSE	P35747 equus cabal
13	2469	69.2	608	2	Q5EG48_MICFO	Q5eg48 microtus fo
14	2462	69.0	608	1	ALBU_RABIT	P49065 oryctolagus
15	2460	68.9	608	2	Q5EG49_MICFO	Q5eg49 microtus fo
16	2455.5	68.8	607	1	ALBU_BOVIN	P02769 bos taurus
17	2438	68.3	608	2	Q5U3X3_RAT	Q5u3x3 rattus norv
18	2437.5	68.3	607	1	ALBU_SHEEP	P14639 ovis aries
19	2431	68.1	608	1	ALBU_RAT	P02770 rattus norv
20	2409.5	67.5	607	1	ALBU_PIG	P08835 sus scrofa
21	2392	67.0	608	2	Q6WDN9_CAVPO	Q6wdn9 cavia porce
22	2387.5	66.9	609	1	ALBU_MERUN	O35090 meriones un
23	2383	66.8	608	1	ALBU_MOUSE	P07724 mus musculu
24	2383	66.8	608	2	Q546G4_MOUSE	Q546g4 mus musculu
25	2379.5	66.7	583	2	Q6B3Z0_ELEMA	Q6b3z0 elephas max
26	2379	66.7	608	2	Q8C7H3_MOUSE	Q8c7h3 mus musculu
27	2336	65.5	576	2	Q8C7C7_MOUSE	Q8c7c7 mus musculu
28	1991	55.8	417	2	Q86YG0_HUMAN	Q86yg0 homo sapien
29	1870.5	52.4	396	2	Q8IUK7_HUMAN	Q8iuk7 homo sapien
30	1562	43.8	615	1	ALBU_CHICK	P19121 gallus gall
31	1295.5	36.3	527	2	Q8JIA9_SPHPU	Q8jia9 sphenodon p
32	1260.5	35.3	609	1	FETA_PANTR	Q28789 pan troglod
33	1256.5	35.2	609	1	FETA_HUMAN	P02771 homo sapien
34	1249.5	35.0	609	1	FETA_GORGO	P28050 gorilla gor
35	1242	34.8	609	2	Q8MJU5_CANFA	Q8mju5 canis famil
36	1242	34.8	626	2	Q8UW05_AMBMC	Q8uw05 ambystoma m
37	1218.5	34.2	610	2	Q8MJ76_PIG	Q8mj76 sus scrofa
38	1215.5	34.1	609	2	Q5CZZ1_XENTR	Q5czz1 xenopus tro
39	1207.5	33.8	607	1	ALBU2_XENLA	P14872 xenopus lae
40	1201.5	33.7	607	2	Q642P7_XENLA	Q642p7 xenopus lae
41	1200	33.6	609	1	FETA_HORSE	P49066 equus cabal
42	1183.5	33.2	608	2	Q7TSF3_MARMO	Q7tsf3 marmota mon
43	1178.5	33.0	615	1	FETA_CHICK	P84407 gallus gall
44	1164.5	32.6	606	1	ALBU1_XENLA	P08759 xenopus lae
45	1087	30.5	624	2	Q8UW06_AMBTE	Q8uw06 ambystoma t

ALIGNMENTS

RESULT 1
ALBU_HUMAN
ID ALBU_HUMAN STANDARD; PRT; 609 AA.
AC P02768; O95574; P04277; Q13140; Q6UXK4; Q9P157; Q9P1I7; Q9UHS3;
AC Q9UJZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Serum albumin precursor.
GN Name=ALB;
GN ORFNames=PRO0903, PRO1708, PRO2044, PRO2619, PRO2675, UNQ696/PRO1341;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczyk A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT LYS-420.
RX MEDLINE=82081882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT GLY-121.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczyk A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RT "Human serum albumin.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT HIROSHIMA-1 LYS-378.
RA Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Fetal liver;
RX MEDLINE=21376145; PubMed=11483580; DOI=10.1101/gr.175501;

RA Yu Y., Zhang C., Zhou G., Wu S., Qu X., Wei H., Xing G., Dong C.,
RA Zhai Y., Wan J., Ouyang S., Li L., Zhang S., Zhou K., Zhang Y., Wu C.,
RA He F.;
RT "Gene expression profiling in human fetal liver and identification of
RT tissue- and developmental-stage-specific genes through compiled
RT expression profiles and efficient cloning of full-length cDNAs.";
RL Genome Res. 11:1392-1403(2001).
RN [7]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP PROTEIN SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573; DOI=10.1016/0014-5793(75)80242-0;
RA Meloun B., Moravek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RP PROTEIN SEQUENCE OF 25-609.
RA Brown J.R., Shockley P., Behrens P.Q.;
RL (In) Bing D.H. (eds.);
RL The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RP NUCLEOTIDE SEQUENCE OF 1-455.
RC TISSUE=Liver;
RA Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
RN [11]
RP NUCLEOTIDE SEQUENCE OF 1-26.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 1-167.
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R.L., Watanabe C., Wieand D., Woods K., Xie M.-H.,
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
RN [13]
RP PROTEIN SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075; DOI=10.1016/0014-5793(76)80496-6;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [14]
RP PROTEIN SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994.";
RL Electrophoresis 15:1459-1465(1994).
RN [15]
RP PROTEIN SEQUENCE OF 166-174.
RX MEDLINE=86242180; PubMed=3087352;
RA Mogard M.H., Kobayashi R., Chen C.F., Lee T.D., Reeve J.R. Jr.,
RA Shively J.E., Walsh J.H.;
RT "The amino acid sequence of kinetensin, a novel peptide isolated from
RT pepsin-treated human plasma: homology with human serum albumin,
RT neurotensin and angiotensin.";
RL Biochem. Biophys. Res. Commun. 136:983-988(1986).
RN [16]
RP PROTEIN SEQUENCE OF 166-174.
RX MEDLINE=87194805; PubMed=2437111;
RA Carraway R.E., Mitra S.P., Cochrane D.E.;
RT "Structure of a biologically active neurotensin-related peptide
RT obtained from pepsin-treated albumin(s).";
RL J. Biol. Chem. 262:5968-5973(1987).
RN [17]
RP DISULFIDE BONDS.
RA Saber M.A., Stockbauer P., Moravek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [18]
RP BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin.";
RL Biochem. J. 171:453-459(1978).
RN [19]
RP VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358; DOI=10.1016/0167-4838(87)90088-4;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
RT domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [20]
RP VARIANTS NASKAPI/MERSIN GLU-396 AND MEXICO GLY-574.
RX MEDLINE=87260818; PubMed=3474609;
RA Takahashi N., Takahashi Y., Blumberg B.S., Putnam F.W.;
RT "Amino acid substitutions in genetic variants of human serum albumin
RT and in sequences inferred from molecular cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4413-4417(1987).
RN [21]
RP VARIANTS NAGASAKI-3 GLN-27 YANOMAMA-2 GLU-396; NAGASAKI-2 ASN-399 AND
RP MAKU GLU-565.
RX MEDLINE=88068523; PubMed=3479777;
RA Takshashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
RT Amerindian and Japanese populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [22]
RP VARIANTS FUKUOKA-2 HIS-23; CHRISTCHURCH/HONOLULU-2 GLN-24; TAGLIACOZZO
RP ASN-337 AND ALBUMIN B/OSAKA-2/PHNOM PHEN LYS-594.
RX MEDLINE=89098947; PubMed=2911589;
RA Arai K., Ishioka N., Huss K., Madison J., Putnam F.W.;
RT "Identical structural changes in inherited albumin variants from
RT different populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:434-438(1989).
RN [23]

```
RP   VARIANTS HONOLULU-2 GLN-24; NAGASAKI-1 GLY-293; HIROSHIMA-1 LYS-378;
RP   TOCHIGI LYS-400; HIROSHIMA-2 LYS-406 AND OSAKA-2 LYS-594.

  Query Match             87.1%;  Score 3108;  DB 1;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-190;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 2
Q645G4_HUMAN
ID   Q645G4_HUMAN PRELIMINARY;      PRT;   609 AA.
AC   Q645G4;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Serum albumin.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Yu Z., Fu Y.;
RT   "High Expression HSA in Pichia for Pharmaceutical Use.";
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY728024; AAU21642.1; -; mRNA.
SQ   SEQUENCE   609 AA;  69366 MW;  F88FF61DD242E818 CRC64;

  Query Match             87.1%;  Score 3108;  DB 2;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-190;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 3
Q5NVH5_PONPY
ID   Q5NVH5_PONPY PRELIMINARY;      PRT;   609 AA.
AC   Q5NVH5;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp459F2310.
GN   Name=DKFZp459F2310;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Cortex;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR926060; CAI29688.1; -; mRNA.
DR   SMR; Q5NVH5; 26-608.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
```

```
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Hypothetical protein.
SQ   SEQUENCE   609 AA;  69366 MW;  F88FF61DD242E818 CRC64;

  Query Match             87.1%;  Score 3108;  DB 2;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-190;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 4
Q68DN5_HUMAN
ID   Q68DN5_HUMAN PRELIMINARY;      PRT;   609 AA.
AC   Q68DN5;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein DKFZp779N1935.
GN   Name=DKFZp779N1935;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RG   The German cDNA Consortium;
RA   Bloecker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA   Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR749331; CAH18185.1; -; mRNA.
DR   SMR; Q68DN5; 26-608.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Hypothetical protein.
SQ   SEQUENCE   609 AA;  69402 MW;  3BA3AFF17BF99E94 CRC64;

  Query Match             86.4%;  Score 3084;  DB 2;  Length 609;
  Best Local Similarity   99.1%;  Pred. No. 2.3e-188;
  Matches  581;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              |||:|||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         24 RDAYKSEVAHRFKDLGEENFKALVLIAFAQYPQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHERTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              |||||||||||||||||||||||||||||||||||||||||| |||||||:|||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQAALVELVKHRPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 5
Q56G89_HUMAN
ID   Q56G89_HUMAN PRELIMINARY;      PRT;   609 AA.
AC   Q56G89;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Serum albumin.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Li H., Zhang Y., Li X., Yang R., Tang S., Zhang M., Hua S.;
RT   "Homo sapiens serum albumin (HSA) cDNA sequence.";
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY960291; AAX63425.1; -; mRNA.
SQ   SEQUENCE   609 AA;  69084 MW;  39B0ECBB1217A99C CRC64;

  Query Match             86.1%;  Score 3073;  DB 2;  Length 609;
  Best Local Similarity   99.1%;  Pred. No. 1.1e-187;
  Matches  581;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQGLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVGSKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              |||||||||||||||||||||||||||||| ||| ||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDCLSVFLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNGRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 6
Q5D0D7_HUMAN
ID   Q5D0D7_HUMAN PRELIMINARY;      PRT;   627 AA.
AC   Q5D0D7;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   ALB protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC039235; AAH39235.1; -; mRNA.
SQ   SEQUENCE   627 AA;  71704 MW;  271C97408D7EDD04 CRC64;

  Query Match             85.9%;  Score 3066;  DB 2;  Length 627;
  Best Local Similarity   98.5%;  Pred. No. 3.3e-187;
  Matches  577;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||      : | |
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKTCCCKSSCLRL 609


RESULT 7
ALBU_MACMU
ID   ALBU_MACMU     STANDARD;      PRT;   600 AA.
```

```
AC   Q28522;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor (Fragment).
GN   Name=ALB;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93211971; PubMed=8460152;
RA   Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA   Dwulet J., Putnam F.W.;
RT   "cDNA and protein sequence of polymorphic macaque albumins that differ
RT   in bilirubin binding.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; M90463; AAA36906.1; -; mRNA.
DR   PIR; A47391; A47391.
DR   HSSP; P02768; 1E7B.
DR   SMR; Q28522; 19-600.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL       <1     10       By similarity.
FT   PROPEP        11     16       By similarity.
FT   CHAIN         17    600       Serum albumin.
FT   DOMAIN        17    197       Albumin 1.
FT   DOMAIN       204    389       Albumin 2.
FT   DOMAIN       396    587       Albumin 3.
FT   METAL         19     19       Copper (By similarity).
FT   BINDING      256    256       Bilirubin (Potential).
FT   DISULFID      69     78       By similarity.
FT   DISULFID      91    107       By similarity.
FT   DISULFID     106    117       By similarity.
FT   DISULFID     140    185       By similarity.
FT   DISULFID     184    193       By similarity.
FT   DISULFID     216    262       By similarity.
FT   DISULFID     261    269       By similarity.
FT   DISULFID     281    295       By similarity.
FT   DISULFID     294    305       By similarity.
FT   DISULFID     332    377       By similarity.
FT   DISULFID     376    385       By similarity.
FT   DISULFID     408    454       By similarity.
FT   DISULFID     453    464       By similarity.
FT   DISULFID     477    493       By similarity.
FT   DISULFID     492    503       By similarity.
FT   DISULFID     530    575       By similarity.
FT   DISULFID     574    583       By similarity.
FT   NON_TER        1      1
```

```
SQ   SEQUENCE   600 AA;  67881 MW;  E45C871A670E740B CRC64;

  Query Match             82.6%;  Score 2947;  DB 1;  Length 600;
  Best Local Similarity   93.5%;  Pred. No. 1.2e-179;
  Matches  546;  Conservative   23;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              || |||||||||||||||:|| ||| :|| :|||||||||:||||||||||||||||||||
Db         16 RDTHKSEVAHRFKDLGEEHFKGLVLVAFSQYLQQCPFEEHVKLVNEVTEFAKTCVADESA 75

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db         76 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPPLVRPE 135

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              |||||||||||| |||||||||:|||||||||||||||| |||||| ||||||||||||
Db        136 VDVMCTAFHDNEATFLKKYLYEVARRHPYFYAPELLFFAARYKAAFAECCQAADKAACLL 195

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              |||||||||||||||||||||||||:||||||||||||:||||||||||||||||||||
Db        196 PKLDELRDEGKASSAKQRLKCASLQKFGDRAFKAWAVARLSQKFPKAEFAEVSKLVTDLT 255

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              |||||||||||||||||||||||:|||||||||||||:|||||||||:|||||||||
Db        256 KVHTECCHGDLLECADDRADLAKYMCENQDSISSKLKECCDKPLLEKSHCLAEVENDEMP 315

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              |||||||||:|||||||||||||||||||||||||||||||||||||:|||||| || ||||
Db        316 ADLPSLAADYVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVMLLLRLAKAYEATLEK 375

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              |||||||||||||||||:|||||||||:|||||||||||||||||||||||||||||||
Db        376 CCAAADPHECYAKVFDEFQPLVEEPQNLVKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 435

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              |||||||||||||:|||| |||||||||||||||||:|||||||||::|||||||
Db        436 TPTLVEVSRNLGKVGAKCCKLPEAKRMPCAEDYLSVVLNRLCVLHEKTPVSEKVTKCCTE 495

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              |||||||||||||:|| |||| |||||||||||:|||||||:|:||||||||||||||||
Db        496 SLVNRRPCFSALELDEAYVPKAFNAETFTFHADMCTLSEKEKQVKKQTALVELVKHKPKA 555

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 672
              |||||| |||:||||||||||||||| |||||| | ||||||||
Db        556 TKEQLKGVMDNFAAFVEKCCKADDKEACFAEEGPKFVAASQAAL 599
```

```
RESULT 8
ALBU_FELCA
ID   ALBU_FELCA     STANDARD;      PRT;   608 AA.
AC   P49064; Q7YSG3;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor (Allergen Fel d 2).
GN   Name=ALB;
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
OC   Felinae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=96194824; PubMed=8647469; DOI=10.1016/0378-1119(95)00851-9;
RA   Hilger C., Grigioni F., Kohnen M., Hentges F.;
RT   "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL   Gene 169:295-296(1996).
RN   [2]
RP   NUCLEOTIDE SEQUENCE OF 25-608.
RC   TISSUE=Liver;
RA   Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA   Rumpold H., Valenta R., Spitzauer S.;
RT   "Escherichia coli expression and purification of recombinant cat
```

```
RT   albumin:IgE recognition, induction of basophil activation and
RT   lymphoproliferative responses in atopic patients.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; X84842; CAA59279.1; -; mRNA.
DR   EMBL; AJ487677; CAD32275.1; -; mRNA.
DR   PIR; JC4660; S57632.
DR   HSSP; P02768; 1E7B.
DR   SMR; P49064; 26-608.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL       1     18       By similarity.
FT   PROPEP      19     24       By similarity.
FT   CHAIN       25    608       Serum albumin.
FT   DOMAIN      25    205       Albumin 1.
FT   DOMAIN     212    397       Albumin 2.
FT   DOMAIN     404    595       Albumin 3.
FT   METAL       27     27       Copper.
FT   DISULFID    77     86       By similarity.
FT   DISULFID    99    115       By similarity.
FT   DISULFID   114    125       By similarity.
FT   DISULFID   148    193       By similarity.
FT   DISULFID   192    201       By similarity.
FT   DISULFID   224    270       By similarity.
FT   DISULFID   269    277       By similarity.
FT   DISULFID   289    303       By similarity.
FT   DISULFID   302    313       By similarity.
FT   DISULFID   340    385       By similarity.
FT   DISULFID   384    393       By similarity.
FT   DISULFID   416    462       By similarity.
FT   DISULFID   461    472       By similarity.
FT   DISULFID   485    501       By similarity.
FT   DISULFID   500    511       By similarity.
FT   DISULFID   538    583       By similarity.
FT   DISULFID   582    591       By similarity.
FT   CONFLICT    75     75       K -> N (in Ref. 2).
FT   CONFLICT    94     94       L -> F (in Ref. 2).
FT   CONFLICT   186    186       K -> R (in Ref. 2).
FT   CONFLICT   251    251       E -> D (in Ref. 2).
FT   CONFLICT   282    282       A -> E (in Ref. 2).
FT   CONFLICT   331    331       V -> A (in Ref. 2).
SQ   SEQUENCE   608 AA;  68659 MW;  07E629CAC5F60E5F CRC64;

  Query Match             73.6%;  Score 2627;  DB 1;  Length 608;
  Best Local Similarity   80.1%;  Pred. No. 3.2e-159;
  Matches  483;  Conservative   53;  Mismatches   57;  Indels   10;  Gaps    1;

Qy         70 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 129
              |:|  |           |  |:|||:||:|||| |||||:|: |||:||:|||||||||||
Db         15 SAYSRG----------VTRREAHQSEIAHRFNDLGEEHFRGLVLVAFSQYLQQCPFEDHV 64

Qy        130 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 189
              |||||||||| ||||:|| ||:|||| | ||||||||:||: |||||||| |:||||||
Db         65 KLVNEVTEFAKGCVADQSAANCEKSLHELLGDKLCTVASLRDKYGEMADCCEKKEPERNE 124

Qy        190 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 249
              |||||||||   :|| || | ||||||:||: || ||||||||||||||||||::|:
Db        125 CFLQHKDDNPGFGQLVTPEADAMCTAFHENEQRFLGKYLYEIARRHPYFYAPELLYYAEE 184

Qy        250 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 309
              ||  |||||:|||||||| ||:| ||::  |||||:|||||||||||||||:|||||
Db        185 YKGVFTECCEAADKAACLTPKVDALREKVLASSAKERLKCASLQKFGERAFKAWSVARLS 244

Qy        310 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 369
              |:|||||||:|||||| |:| ||||||||||||||||||||||||||||:||||||
Db        245 QKFPKAEFAEISKLVTDLAKIHKECCHGDLLECADDRADLAKYICENQDSISTKLKECCG 304

Qy        370 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 429
              ||:|||||||:||| ||:||||| || |||| |:||||| |||||||| |||||:||||:
Db        305 KPVLEKSHCISEVERDELPADLPPLAVDFVEDKEVCKNYQEAKDVFLGTFLYEYSRRHPE 364

Qy        430 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 489
              ||| ||||||| || |||||||  ||  ||| ||||||||||| ||:| ||||||:|||
Db        365 YSVSLLLRLAKEYEATLEKCCATDDPPACYAHVFDEFKPLVEEPHNLVKTNCELFEKLGE 424

Qy        490 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 549
              | |||||||||||||||||||||||:|||||||| ||||:|: ||||||||||:|
Db        425 YGFQNALLVRYTKKVPQVSTPTLVEVSRSLGKVGSKCCTHPEAERLSCAEDYLSVVLNRL 484

Qy        550 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 609
              |||||||||:|||||||||||||||||:||||||||:|||||||||:|| | |
Db        485 CVLHEKTPVSERVTKCCTESLVNRRPCFSALQVDETYVPKEFSAETFTFHADLCTLPEAE 544

Qy        610 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 669
              :|||||:||||||:|||||:|||| || || :||:||| |:||| |||||| |||||:|
Db        545 KQIKKQSALVELLKHKPKATEEQLKTVMGDFGSFVDKCCAAEDKEACFAEEGPKLVAAAQ 604

Qy        670 AAL 672
              |||
Db        605 AAL 607


RESULT 9
ALBU_CANFA
ID   ALBU_CANFA     STANDARD;      PRT;   608 AA.
AC   P49822; O77705; Q9TSZ4;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor (Allergen Can f 3).
GN   Name=ALB;
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC   Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA].
RC   STRAIN=Beagle; TISSUE=Liver;
RA   Hilger C.;
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=20148667; PubMed=10669848; DOI=10.1016/S0091-6749(00)90077-0;
RA   Pandjaitan B., Swoboda I., Brandejsky-Pichler F., Rumpold H.,
RA   Valenta R., Spitzauer S.;
RT   "Escherichia coli expression and purification of recombinant dog
RT   albumin, a cross-reactive animal allergen.";
RL   J. Allergy Clin. Immunol. 105:279-285(2000).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [MRNA].
RC   STRAIN=Beagle; TISSUE=Liver;
RA   Miyake M., Okazaki M., Iwabuchi S.;
```

```
RT   "Isolation of a cDNA encoding canine serum albumin.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   PROTEIN SEQUENCE OF 25-48.
RX   MEDLINE=75011422; PubMed=4414013;
RA   Dixon J.W., Sarkar B.;
RT   "Isolation, amino acid sequence and copper(II)-binding properties of
RT   peptide (1-24) of dog serum albumin.";
RL   J. Biol. Chem. 249:5872-5877(1974).
RN   [5]
RP   PROTEIN SEQUENCE OF 25-38.
RC   TISSUE=Heart;
RX   MEDLINE=98163340; PubMed=9504812;
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
RN   [6]
RP   NUCLEOTIDE SEQUENCE OF 215-478.
RC   TISSUE=Salivary gland;
RX   MEDLINE=94201492; PubMed=7512102;
RA   Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,
RA   Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
RT   "Molecular characterization of dog albumin as a cross-reactive
RT   allergen.";
RL   J. Allergy Clin. Immunol. 93:614-627(1994).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AJ133489; CAB64867.1; -; mRNA.
DR   EMBL; Y17737; CAA76841.1; -; mRNA.
DR   EMBL; AB090854; BAC10663.1; -; mRNA.
DR   EMBL; S72946; AAB30434.1; -; mRNA.
DR   HSSP; P02768; 1E7B.
DR   SMR; P49822; 26-607.
DR   HSC-2DPAGE; P49822; DOG.
DR   Ensembl; ENSCAFG00000003016; Canis familiaris.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Allergen; Copper; Direct protein sequencing; Lipid-binding;
KW   Metal-binding; Repeat; Signal.
FT   SIGNAL        1     18       Potential.
FT   PROPEP       19     24
FT   CHAIN        25    608       Serum albumin.
FT   DOMAIN       25    205       Albumin 1.
FT   DOMAIN      212    397       Albumin 2.
FT   DOMAIN      404    595       Albumin 3.
FT   METAL        27     27       Copper (By similarity).
FT   DISULFID     77     86       By similarity.
FT   DISULFID     99    115       By similarity.
FT   DISULFID    114    125       By similarity.
FT   DISULFID    148    193       By similarity.
FT   DISULFID    192    201       By similarity.
FT   DISULFID    224    270       By similarity.
FT   DISULFID    269    277       By similarity.
FT   DISULFID    289    303       By similarity.
FT   DISULFID    302    313       By similarity.
FT   DISULFID    340    385       By similarity.
FT   DISULFID    384    393       By similarity.
FT   DISULFID    416    462       By similarity.
FT   DISULFID    461    472       By similarity.
FT   DISULFID    485    501       By similarity.
FT   DISULFID    500    511       By similarity.
FT   DISULFID    538    583       By similarity.
FT   DISULFID    582    591       By similarity.
FT   CONFLICT      1     26       MKWVTFISLFFLFSSAYSRGLVRREA -> MDT (in
FT                                Ref. 2).
FT   CONFLICT    146    146       A -> R (in Ref. 2).
FT   CONFLICT    206    206       I -> T (in Ref. 2).
FT   CONFLICT    349    349       V -> A (in Ref. 2).
FT   CONFLICT    359    359       A -> S (in Ref. 2 and 6).
FT   CONFLICT    448    448       V -> VV (in Ref. 6).
FT   CONFLICT    474    474       E -> D (in Ref. 2 and 6).
SQ   SEQUENCE   608 AA;  68605 MW;  3DB012FF7C979CF3 CRC64;

  Query Match             72.1%;  Score 2574;  DB 1;  Length 608;
  Best Local Similarity   78.4%;  Pred. No. 7.8e-156;
  Matches  473;  Conservative   57;  Mismatches   63;  Indels   10;  Gaps    2;

Qy         70 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 129
              |:|  |           ||: |:|:|||:||: |||||:|: |||:||:|||||||||||
Db         15 SAYSRG---------LVR-REAYKSEIAHRYNDLGEEHFRGLVLVAFSQYLQQCPFEDHV 64

Qy        130 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 189
              ||  ||||||| | |:|| ||||||||||||||||||:||: ||:||||| ||||:|||
Db         65 KLAKEVTEFAKACAAEESGANCDKSLHTLFGDKLCTVASLRDKYGDMADCCEKQEPDRNE 124

Qy        190 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 249
              ||| ||||||  | || || | :| || |||: || ||||||||||||||||||::|::
Db        125 CFLAHKDDNPGFPPLVAPEPDALCAAFQDNEQLFLGKYLYEIARRHPYFYAPELLYYAQQ 184

Qy        250 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 309
              ||  | ||||||||||| ||:: ||::   ||||:| |||||||||:||||||:|||||
Db        185 YKGVFAECCQAADKAACLGPKIEALREKVLLSSAKERFKCASLQKFGDRAFKAWSVARLS 244

Qy        310 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 369
              ||||||:|||:||:|||||||| ||||||||||||||||||:||||||||:||||||:
Db        245 QRFPKADFAEISKVVTDLTKVHKECCHGDLLECADDRADLAKYMCENQDSISTKLKECCD 304

Qy        370 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 429
              ||:|||| |:|||| ||:| |||||||||| |:||||| |||||||| ||||||||||:
Db        305 KPVLEKSQCLAEVERDELPGDLPSLAADFVEDKEVCKNYQEAKDVFLGTFLYEYARRHPE 364

Qy        430 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 489
              ||| ||||||| || |||||||  ||  ||||| |||||||:|||||:| ||||||:|||
Db        365 YSVSLLLRLAKEYEATLEKCCATDDPPTCYAKVLDEFKPLVDEPQNLVKTNCELFEKLGE 424

Qy        490 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 549
              | |||||||||||| |||||||||||| |||||:|||| ||::|| |||||:||||||:|
Db        425 YGFQNALLVRYTKKAPQVSTPTLVEVSRKLGKVGTKCCKKPESERMSCAEDFLSVVLNRL 484

Qy        550 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 609
              ||||||||||:||||||:|||||||||||| ||||||||||||||||||||:||| | |
Db        485 CVLHEKTPVSERVTKCCSESLVNRRPCFSGLEVDETYVPKEFNAETFTFHADLCTLPEAE 544

Qy        610 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 669
              :|:|||||||:|||||| |||| || || ||||||| |::|| ||:||| |||||:|
Db        545 KQVKKQTALVELLKHKPKATDEQLKTVMGDFGAFVEKCCAAENKEGCFSEEGPKLVAAAQ 604

Qy        670 AAL 672
              |||
Db        605 AAL 607


RESULT 10
Q95VB7_SCHMA
ID   Q95VB7_SCHMA PRELIMINARY;      PRT;   608 AA.
AC   Q95VB7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Albumin.
OS   Schistosoma mansoni (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6183;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF418550; AAL08579.1; -; mRNA.
DR   HSSP; P02768; 1HK1.
DR   SMR; Q95VB7; 26-608.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0008289; F:lipid binding; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 2.
SQ   SEQUENCE   608 AA;  68225 MW;  E5EABB28E1C66E54 CRC64;

  Query Match           70.3%;  Score 2509;  DB 2;  Length 608;
  Best Local Similarity  76.4%;  Pred. No. 1.1e-151;
  Matches  446;  Conservative   79;  Mismatches   59;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              |||||||:|||||||||::|| ||||||:|:||:||:|:|||||||||:||||||||||
Db         24 RDAHKSEIAHRFKDLGEQHFKGLVLIAFSQFLQKCPYEEHVKLVNEVTDFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              |||||||||||||||| : |||::|||:||||||:|||||||||:||||:||||  |||:
Db         84 ENCDKSLHTLFGDKLCAIPTLRDSYGELADCCAKKEPERNECFLKHKDDHPNLPPFVRPD 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
               : |||:| :|  ||:  ||:|:||||||||||||::|::| |  ||||  ||||||:
Db        144 AEAMCTSFQENAVTFMGHYLHEVARRHPYFYAPELLYYAEKYSAIMTECCGEADKAACIT 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              |||| |:::  |||  |||||:|||:||:||||||||||:||:||||:|||::|| ||||
Db        204 PKLDALKEKALASSVNQRLKCSSLQRFGQRAFKAWAVARMSQKFPKADFAEITKLATDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              |:  |||||||||||||||:||||:|||| |||||: ||:||:|:||||::|||||::|
Db        264 KLTEECCHGDLLECADDRAELAKYMCENQASISSKLQACCDKPVLKKSHCLSEVENDDLP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              |||||||||||| |:|||||||||||||| |||||||||||||| ||||||| || ||||
Db        324 ADLPSLAADFVEDKEVCKNYAEAKDVFLGTFLYEYARRHPDYSVALLLRLAKKYEATLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              |||  |||   || || |||:||||||:||:| ||||||:|||| |||||:||||:| ||||
Db        384 CCAEADPSACYGKVLDEFQPLVEEPKNLVKANCELFEKLGEYGFQNALIVRYTQKAPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||  :|||||||||||   ||||:|:|| |||:|  :||::||||||||::||||||
Db        444 TPTLVEAARNLGKVGSKCCVLPEAQRLPCVEDYISAILNRVCVLHEKTPVSEQVTKCCTG 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              |:| |||||||| |||||||||| ||||||||||:| |||:|:||| ||||||||||||
Db        504 SVVERRPCFSALPVDETYVPKEFKAETFTFHADICSLPEKEKQMKKQAALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 672
              |  ||: |: :| ||::|||||:||| ||:|:| ||||:|||||
Db        564 TGPQLRTVLGEFTAFLDKCCKAEDKEACFSEDGPKLVASSQAAL 607

RESULT 11
```

```
ALBU_EQUAS
ID   ALBU_EQUAS     STANDARD;      PRT;   607 AA.
AC   Q5XLE4;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor.
GN   Name=ALB;
OS   Equus asinus (Donkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9793;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Li H., Tang Y., Pingfan R.;
RT   "Full-length cDNA sequence of serum albumin of donkey (Equus asinus)
RT   and its structure analysis.";
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AY754333; AAV28861.1; -; mRNA.
DR   SMR; Q5XLE4; 27-607.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL        1     18       By similarity.
FT   PROPEP       19     24       By similarity.
FT   CHAIN        25    607       Serum albumin.
FT   DOMAIN       25    204       Albumin 1.
FT   DOMAIN      211    396       Albumin 2.
FT   DOMAIN      403    594       Albumin 3.
FT   METAL        27     27       Copper (By similarity).
FT   DISULFID     77     86       By similarity.
FT   DISULFID     99    115       By similarity.
FT   DISULFID    114    125       By similarity.
FT   DISULFID    147    192       By similarity.
FT   DISULFID    191    200       By similarity.
FT   DISULFID    223    269       By similarity.
FT   DISULFID    268    276       By similarity.
FT   DISULFID    288    302       By similarity.
FT   DISULFID    301    312       By similarity.
FT   DISULFID    339    384       By similarity.
FT   DISULFID    383    392       By similarity.
FT   DISULFID    415    461       By similarity.
FT   DISULFID    460    471       By similarity.
FT   DISULFID    484    500       By similarity.
FT   DISULFID    499    510       By similarity.
FT   DISULFID    537    582       By similarity.
FT   DISULFID    581    590       By similarity.
SQ   SEQUENCE   607 AA;  68539 MW;  7099E1E08E3C426A CRC64;

  Query Match           70.1%;  Score 2501.5;  DB 1;  Length 607;
  Best Local Similarity  75.0%;  Pred. No. 3.3e-151;
  Matches  452;  Conservative   71;  Mismatches   69;  Indels   11;  Gaps    2;
```

```
Qy         70 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 129
              |:|  |           |  || ||||:|||| ||||::|| |||:||:|||||||||||||
Db         15 SAYFRG----------VLRRDTHKSEIAHRFNDLGEKHFKGLVLVAFSQYLQQCPFEDHV 64

Qy        130 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 189
              |||||||||| | ||||||||||||||||||||||||| ||||:|||| ||||||||
Db         65 KLVNEVTEFAKKCAADESAENCDKSLHTLFGDKLCTVATLRATYGELADCCEKQEPERNE 124

Qy        190 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 249
              ||| ||||:||||:| :|| |  | || :: : || |||||:|||||||| ||||| |:
Db        125 CFLTHKDDHPNLPKL-KPEPDAQCAAFQEDPDKFLGKYLYEVARRHPYFYGPELLFHAEE 183

Qy        250 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 309
              ||| ||||| | ||| ||:|||| |::    ||||:||||:| ||||||||||:|||||
Db        184 YKADFTECCPADDKAGCLIPKLDALKERILLSSAKERLKCSSFQKFGERAFKAWSVARLS 243

Qy        310 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 369
              |:||||:|||||:|||||||| |||||||||||||||| |||||:||||| ||| ||:
Db        244 QKFPKADFAEVSKIVTDLTKVHKECCHGDLLECADDRADLTKYICEHQDSISGKLKACCD 303

Qy        370 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 429
              ||||:||||||||: |::|:|||:||||| | |::||:| :||||||| |||||:|||||
Db        304 KPLLQKSHCIAEVKEDDLPSDLPALAADFAEDKEICKHYKDAKDVFLGTFLYEYSRRHPD 363

Qy        430 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 489
              ||| ||||:||||| ||||||| |||  ||| |||:| ||||||::|:|:||:|||:||
Db        364 YSVSLLLRIAKTYEATLEKCCAEADPPACYATVFDQFTPLVEEPKSLVKKNCDLFEEVGE 423

Qy        490 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 549
              | |||||:|||||| |||||||||||: | ||||||:||| ||::|:||:|::|:: ||:|
Db        424 YDFQNALIVRYTKKAPQVSTPTLVEIGRTLGKVGSRCCKLPESERLPCSENHLALALNRL 483

Qy        550 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 609
              ||||||||||:::|||||:||  |||||||||:|| |:|||| |||||||||||| | |
Db        484 CVLHEKTPVSEKITKCCTDSLAERRPCFSALELDEGYIPKEFKAETFTFHADICTLPEDE 543

Qy        610 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 669
              :|||||:|| |||||||||||||| |: :|:||| ||| |:||| |||||| ||||:||
Db        544 KQIKKQSALAELVKHKPKATKEQLKTVLGNFSAFVAKCCGAEDKEACFAEEGPKLVASSQ 603

Qy        670 AAL 672
               ||
Db        604 LAL 606
```

```
RESULT 12
ALBU_HORSE
ID   ALBU_HORSE     STANDARD;      PRT;   607 AA.
AC   P35747;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor (Allergen Equ c 3).
GN   Name=ALB;
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC   TISSUE=Liver;
RX   MEDLINE=93345495; PubMed=8344282;
RA   Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT   "X-ray and primary structure of horse serum albumin (Equus caballus)
RT   at 0.27-nm resolution.";
RL   Eur. J. Biochem. 215:205-212(1993).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human. Binds to IgE.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; X74045; CAA52194.1; -; mRNA.
DR   PIR; S34053; ABHOS.
DR   HSSP; P02768; 1HK1.
DR   SMR; P35747; 27-607.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL       1     18       By similarity.
FT   PROPEP      19     24       By similarity.
FT   CHAIN       25    607       Serum albumin.
FT   DOMAIN      25    204       Albumin 1.
FT   DOMAIN     211    396       Albumin 2.
FT   DOMAIN     403    594       Albumin 3.
FT   METAL       27     27       Copper (By similarity).
FT   DISULFID    77     86
FT   DISULFID    99    115
FT   DISULFID   114    125
FT   DISULFID   147    192
FT   DISULFID   191    200
FT   DISULFID   223    269
FT   DISULFID   268    276
FT   DISULFID   288    302
FT   DISULFID   301    312
FT   DISULFID   339    384
FT   DISULFID   383    392
FT   DISULFID   415    461
FT   DISULFID   460    471
FT   DISULFID   484    500
FT   DISULFID   499    510
FT   DISULFID   537    582
FT   DISULFID   581    590
SQ   SEQUENCE   607 AA;  68599 MW;  256F6E830A1B90C5 CRC64;
```

```
  Query Match             69.5%;  Score 2481.5;  DB 1;  Length 607;
  Best Local Similarity   74.6%;  Pred. No. 6.2e-150;
  Matches  450;  Conservative   70;  Mismatches   72;  Indels   11;  Gaps    2;

Qy         70 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 129
              |:|  |           |  || ||||:|||| ||||::|| |||:||:|||||||||||||
Db         15 SAYSRG----------VLRRDTHKSEIAHRFNDLGEKHFKGLVLVAFSQYLQQCPFEDHV 64

Qy        130 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 189
              |||||||||| | ||||||||||||||||||||||||| ||||:|||| ||||||||
Db         65 KLVNEVTEFAKKCAADESAENCDKSLHTLFGDKLCTVATLRATYGELADCCEKQEPERNE 124

Qy        190 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 249
              ||| ||||:||||:| :|| |  | || :: : || |||||:|||||||| ||||| |:
Db        125 CFLTHKDDHPNLPKL-KPEPDAQCAAFQEDPDKFLGKYLYEVARRHPYFYGPELLFHAEE 183

Qy        250 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 309
              ||| ||||| | || |||:|||| |::    ||||:||||:| | ||||| |||:|||||
Db        184 YKADFTECCPADDKLACLIPKLDALKERILLSSAKERLKCSSFQNFGERAVKAWSVARLS 243

Qy        310 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 369
              |:||||:||||||:|||||||| ||||||||||||||||||||:||||| ||| ||:
Db        244 QKFPKADFAEVSKIVTDLTKVHKECCHGDLLECADDRADLAKYICEHQDSISGKLKACCD 303

Qy        370 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 429
              ||||:||||||||: |::|:|||:||||| | |::||:| :|||||||| |||||:|||||
Db        304 KPLLQKSHCIAEVKEDDLPSDLPALAADFAEDKEICKHYKDAKDVFLGTFLYEYSRRHPD 363
```

```
Qy        430 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 489
              ||| ||||:||||| ||||||| ||| || |||:| ||||||::|:|:||:|||::||
Db        364 YSVSLLLRIAKTYEATLEKCCAEADPPACYRTVFDQFTPLVEEPKSLVKKNCDLFEEVGE 423

Qy        490 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 549
              | |||||:|||||| |||||||||||: | ||||||:||| ||::|:||:|::|:: ||:|
Db        424 YDFQNALIVRYTKKAPQVSTPTLVEIGRTLGKVGSRCCKLPESERLPCSENHLALALNRL 483

Qy        550 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 609
              |||||||||||:::|||||:||  |||||||||:|| |||||| |||||||||||||| | |
Db        484 CVLHEKTPVSEKITKCCTDSLAERRPCFSALELDEGYVPKEFKAETFTFHADICTLPEDE 543

Qy        610 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 669
              :|||||:|| |||||||||||||| |: :|:||| |||  :||| |||||| ||||:||
Db        544 KQIKKQSALAELVKHKPKATKEQLKTVLGNFSAFVAKCCGREDKEACFAEEGPKLVASSQ 603

Qy        670 AAL 672
               ||
Db        604 LAL 606
```

RESULT 13
Q5EG48_MICFO
ID Q5EG48_MICFO PRELIMINARY; PRT; 608 AA.
AC Q5EG48;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Albumin.
OS Microtus fortis calamorum.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Arvicolinae; Microtus.
OX NCBI_TaxID=311220;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Hu W.-X., Wu G.-J., Qin Z.-Q., Luo S.-Q.;
RT "Albumin gene of Microtus Fortis calamorum liver.";
RL Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY885265; AAW79113.1; -; mRNA.
DR InterPro; IPR001703; Alphafetoprot.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Serum_albumin; 3.
DR PRINTS; PR00803; AFETOPROTEIN.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68308 MW; B04A061339494D3F CRC64;

  Query Match 69.2%; Score 2469; DB 2; Length 608;
  Best Local Similarity 76.2%; Pred. No. 3.9e-149;
  Matches 445; Conservative 65; Mismatches 74; Indels 0; Gaps 0;

```
Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              |||||||:|||: |||||: || |||| |||:||:||:|:|:|||||||:||| | |||||
Db         24 RDAHKSEIAHRYNDLGEKYFKGLVLITFAQHLQKCPYEEHIKLVNEVTDFAKACAADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              |||||||||||||||  :  | : | |||:||||||||||:|||:|||| ||||  ||||
Db         84 ENCDKSLHTLFGDKLCAIPNLGDNYAEMAECCAKQEPERNQCFLKHKDDKPNLPPFVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
               :||||:| :|  :|:  ||: :|||||||||||::|::  |  ||||  |||||||
Db        144 AEVMCTSFQENPASFMGHYLHAVARRHPYFYAPELLYYAEKQSAIMTECCAEADKAACLG 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              |||| |:::   || ||||||:||:||||||||||||||:||:||||:|||::|| ||||
Db        204 PKLDALKEKALLSSVKQRLKCSSLEKFGERAFKAWAVARMSQKFPKADFAEITKLATDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              || ||||||||||||||| :||||:|:|| :|||||  ||:||:|:|:|:||:|||::||||
Db        264 KVTQECCHGDLLECADDRLELAKYMCDNQATISSKLHTCCDKPVLQKAHCLAEVDHDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              |||  || |||| |||||||||||||||| |||||||||||||| ||||||| || ||||
Db        324 ADLTPLADDFVEDKDVCKNYAEAKDVFLGTFLYEYARRHPDYSVALLLRLAKKYEATLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||| |||| || ||||||:||||||:||:| ||||||:|||| ||||||||||:| ||||
Db        384 CCAEADPHACYGKVFDEFQPLVEEPKNLVKANCELFEKLGEYGFQNALLVRYTQKAPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              |||||||||||||:|||  ||| |:|| ||||| :|| ||||||||||::||||||
Db        444 TPTLVEVSRNLGKVGTKCCALPEADRLPCVEDYLSAILNHLCVLHEKTPVSEQVTKCCTG 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              |:| |||||||| ||| |||||| ||:|||||||||| :||:|||||||| |||||||||
Db        504 SVVERRPCFSALPVDEHYVPKEFKAESFTFHADICTLPDKEKQIKKQTALAELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 672
              | :||| || :|:||:|||||||||| ||:||| |||| |||||
Db        564 TGDQLKTVMGEFSAFLEKCCKADDKEACFSEEGPKLVATSQAAL 607

RESULT 14
ALBU_RABIT
ID  ALBU_RABIT  STANDARD;  PRT;  608 AA.
AC  P49065;
DT  01-FEB-1996 (Rel. 33, Created)
DT  29-MAR-2004 (Rel. 43, Last sequence update)
DT  10-MAY-2005 (Rel. 47, Last annotation update)
DE  Serum albumin precursor.
GN  Name=ALB;
OS  Oryctolagus cuniculus (Rabbit).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC  Oryctolagus.
OX  NCBI_TaxID=9986;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=New Zealand white; TISSUE=Liver;
RX  MEDLINE=97275135; PubMed=9129029;
RA  Syed S., Schuyler P.D., Kulczycky M., Sheffield W.P.;
RT  "Potent antithrombin activity and delayed clearance from the
RT  circulation characterize recombinant hirudin genetically fused to
RT  albumin.";
RL  Blood 89:3243-3252(1997).
RN  [2]
RP  SEQUENCE REVISION TO 322-323 AND 506-507.
RA  Sheffield W.P.;
RL  Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
CC  -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC      binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC      hormones, bilirubin and drugs. Its main function is the regulation
CC      of the colloidal osmotic pressure of blood.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- TISSUE SPECIFICITY: Plasma.
CC  -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC  -!- SIMILARITY: Contains 3 albumin domains.
CC  ----------------------------------------------------------------------
CC  This Swiss-Prot entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use as long as its content is in no way modified and this statement is not
CC  removed.
CC  ----------------------------------------------------------------------
DR  EMBL; U18344; AAB58347.2; -; mRNA.
DR  HSSP; P02768; 1E7B.
DR  SMR; P49065; 26-608.
DR  InterPro; IPR001703; Alphafetoprot.
DR  InterPro; IPR000264; Serum_albumin.
DR  Pfam; PF00273; Serum_albumin; 3.
DR  PRINTS; PR00803; AFETOPROTEIN.
```

```
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL        1     18       By similarity.
FT   PROPEP       19     24       By similarity.
FT   CHAIN        25    608       Serum albumin.
FT   DOMAIN       25    205       Albumin 1.
FT   DOMAIN      212    397       Albumin 2.
FT   DOMAIN      404    595       Albumin 3.
FT   METAL        27     27       Copper.
FT   DISULFID     77     86       By similarity.
FT   DISULFID     99    115       By similarity.
FT   DISULFID    114    125       By similarity.
FT   DISULFID    148    193       By similarity.
FT   DISULFID    192    201       By similarity.
FT   DISULFID    224    270       By similarity.
FT   DISULFID    269    277       By similarity.
FT   DISULFID    289    303       By similarity.
FT   DISULFID    302    313       By similarity.
FT   DISULFID    340    385       By similarity.
FT   DISULFID    384    393       By similarity.
FT   DISULFID    416    462       By similarity.
FT   DISULFID    461    472       By similarity.
FT   DISULFID    485    501       By similarity.
FT   DISULFID    500    511       By similarity.
FT   DISULFID    538    583       By similarity.
FT   DISULFID    582    591       By similarity.
SQ   SEQUENCE   608 AA;  68910 MW;  9EECAFDA86B1EF09 CRC64;

  Query Match             69.0%;  Score 2462;  DB 1;  Length 608;
  Best Local Similarity   74.4%;  Pred. No. 1.1e-148;
  Matches  435;  Conservative   78;  Mismatches   72;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              |:|||||:|||| |:|||:|  |||| |:||||:||:|:| ||| |||: || |||||||
Db         24 REAHKSEIAHRFNDVGEEHFIGLVLITFSQYLQKCPYEEHAKLVKEVTDLAKACVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||| :||||:| : :||:|||::|||| |:|||||||| |||| |:||    |||
Db         84 ANCDKSLHDIFGDKICALPSLRDTYGDVADCCEKKEPERNECFLHHKDDKPDLPPFARPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||:| ||||:|: |   ||||:|||||||||||::|::|||  ||||:|||| |||
Db        144 ADVLCKAFHDDEKAFFGHYLYEVARRHPYFYAPELLYYAQKYKAILTECCEAADKGACLT 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              |||| |  :   |:|::|:|||:||||:||:|||: |||||||||:| ::||:|||||
Db        204 PKLDALEGKSLISAAQERLRCASIQKFGDRAYKAWALVRLSQRFPKADFTDISKIVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||| |||||||||||||||||:||:|::||| |||||:||:|||:|||  : ||| |
Db        264 KVHKECCHGDLLECADDRADLAKYMCEHQETISSHLKECCDKPILEKAHCIYGLHNDETP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              | ||::| :||| ||||||| ||||:||| |||||:||||||||||||| | || ||:|
Db        324 AGLPAVAEEFVEDKDVCKNYEEAKDLFLGKFLYEYSRRHPDYSVVLLLRLGKAYEATLKK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              |||  ||| ||||| |||:|||:||:||:|||||:||||:| ||||||||||||||||
Db        384 CCATDDPHACYAKVLDEFQPLVDEPKNLVKQNCELYEQLGDYNFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||:||:|||||||||||||:|:|| |||||||||:|||||||||::||||||:|
Db        444 TPTLVEISRSLGKVGSKCCKHPEAERLPCVEDYLSVVLNRLCVLHEKTPVSEKVTKCCSE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              |||:||||||||  |||||||||||||||||||||| | ||:|||||||||||||| |
Db        504 SLVDRRPCFSALGPDETYVPKEFNAETFTFHADICTLPETERKIKKQTALVELVKHKPHA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALG 673
              | :||| |: :| | ::||| |:||| ||| || ||| :|:| ||
Db        564 TNDQLKTVVGEFTALLDKCCSAEDKEACFAVEGPKLVESSKATLG 608


RESULT 15
Q5EG49_MICFO
ID   Q5EG49_MICFO PRELIMINARY;      PRT;   608 AA.
AC   Q5EG49;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Albumin.
OS   Microtus fortis fortis.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Arvicolinae; Microtus.
OX   NCBI_TaxID=311338;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Hu W.-X., Wu G.-J., Qin Z.-Q., Luo S.-Q.;
RT   "Albumin gene of Microtus fortis fortis liver.";
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY885264; AAW79112.1; -; mRNA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 2.
SQ   SEQUENCE   608 AA;  68309 MW;  060802BCD5E219D2 CRC64;

  Query Match             68.9%;  Score 2460;  DB 2;  Length 608;
  Best Local Similarity   75.9%;  Pred. No. 1.5e-148;
  Matches  443;  Conservative   66;  Mismatches   75;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              |||||||:|||: |||| || |||| |||:||:||:|:|:|||||||:||| | |||||
Db         24 RDAHKSEIAHRYNDLGEKYFKGLVLITFAQHLQKCPYEEHIKLVNEVTDFAKACAADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              |||||||||||||||:  | : | |||:||||||||||:|||:|||| |:||  ||||
Db         84 ENCDKSLHTLFGDKLCAIPNLGDNYAEMAECCAKQEPERNQCFLKHKDDKPSLPPFVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              :||||:| :|  :|:  ||: :|||||||||||::|::  |  ||||  |||||||
Db        144 AEVMCTSFQENPASFMGHYLHAVARRHPYFYAPELLYYAEKQSAIMTECCAEADKAACLG 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              |||| |:::   || |||||:||:|||||||||||||:||:||||:|||::|| ||||
Db        204 PKLDALKEKALLSSVKQRLKCSSLEKFGERAFKAWAVARMSQKFPKADFAEITKLATDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||  ||||||||||| ||| :||||:|:|| :|||||   ||:||:|:|:||:|||::|||||
Db        264 KVTQECCHGDLLECVDDRLELAKYMCDNQATISSKLHTCCDKPVLQKAHCLAEVDHDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              |||  || |||| |||||||||||||| ||||||||||||| ||||||| || ||||
Db        324 ADLTPLADDFVEDKDVCKNYAEAKDVFLGTFLYEYARRHPDYSVALLLRLAKKYEATLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||| |||| || ||||||:||||||:||:| ||||||:|||| ||||||||||:| ||||
Db        384 CCAEADPHACYGKVFDEFQPLVEEPKNLVKANCELFEKLGEYGFQNALLVRYTQKAPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              |||||||||||||||:|||  ||| |:|| ||||| :|| |||||||||||::||||||
Db        444 TPTLVEVSRNLGKVGTKCCALPEADRLPCVEDYLSAILNHLCVLHEKTPVSEQVTKCCTG 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              |:| |||||||| ||| |||||| ||:|||||||||| :||:|||||||| |||||||||
Db        504 SVVERRPCFSALPVDEHYVPKEFKAESFTFHADICTLPDKEKQIKKQTALAELVKHKPKA 563
```

```
Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 672
              | :||| || :|:||:|||||||||| ||:||| |||| |||||
Db        564 TGDQLKTVMGEFSAFLEKCCKADDKEACFSEEGPKLVATSQAAL 607
```

Search completed: April 19, 2006, 12:08:49
Job time : 181.074 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:09:12 ; Search time 42.6037 Seconds
(without alignments)
1307.948 Million cell updates/sec

Title: US-10-775-180-447
Perfect score: 3568
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....TCFAEEGKKLVAASQAALGL 674

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3112.5	87.2	787	1	US-08-256-938-4	Sequence 4, Appli
2	3112.5	87.2	787	1	US-08-797-689-16	Sequence 16, Appl
3	3112.5	87.2	787	2	US-09-984-186-16	Sequence 16, Appl
4	3108	87.1	609	2	US-09-976-594-977	Sequence 977, App
5	3108	87.1	609	2	US-09-919-039-370	Sequence 370, App
6	3108	87.1	610	1	US-08-797-689-2	Sequence 2, Appli
7	3108	87.1	610	2	US-09-984-186-2	Sequence 2, Appli
8	3108	87.1	622	2	US-09-949-016-11170	Sequence 11170, A
9	3108	87.1	783	1	US-08-256-938-2	Sequence 2, Appli
10	3104	87.0	609	1	US-08-222-619-3	Sequence 3, Appli
11	3104	87.0	609	1	US-08-433-037-4	Sequence 4, Appli
12	3104	87.0	609	2	US-08-897-956A-2	Sequence 2, Appli
13	3104	87.0	609	4	PCT-US95-04075-3	Sequence 3, Appli
14	3103.5	87.0	978	2	US-08-897-956A-3	Sequence 3, Appli
15	3103	87.0	585	1	US-08-153-799-14	Sequence 14, Appl
16	3103	87.0	585	1	US-08-702-572-2	Sequence 2, Appli
17	3103	87.0	585	2	US-08-769-746-2	Sequence 2, Appli
18	3103	87.0	585	2	US-09-833-118A-18	Sequence 18, Appl
19	3103	87.0	585	2	US-09-832-929A-18	Sequence 18, Appl
20	3103	87.0	585	2	US-09-833-111A-18	Sequence 18, Appl
21	3093	86.7	585	1	US-08-448-196A-3	Sequence 3, Appli
22	3093	86.7	585	1	US-08-984-176-1	Sequence 1, Appli
23	2458.5	68.9	583	1	US-08-448-196A-5	Sequence 5, Appli
24	2450.5	68.7	583	1	US-08-448-196A-4	Sequence 4, Appli
25	2450.5	68.7	583	2	US-10-360-101-200	Sequence 200, App
26	2432.5	68.2	583	1	US-08-448-196A-6	Sequence 6, Appli
27	2426	68.0	584	1	US-08-448-196A-7	Sequence 7, Appli
28	2393.5	67.1	604	2	US-10-045-170A-1	Sequence 1, Appli
29	2389	67.0	582	1	US-08-134-638-1	Sequence 1, Appli
30	1256.5	35.2	609	1	US-08-222-619-4	Sequence 4, Appli
31	1256.5	35.2	609	2	US-09-976-594-456	Sequence 456, App
32	1256.5	35.2	609	4	PCT-US95-04075-4	Sequence 4, Appli
33	1256.5	35.2	612	2	US-09-949-016-11201	Sequence 11201, A
34	1213.5	34.0	609	2	US-09-186-949A-2	Sequence 2, Appli
35	1206.5	33.8	590	1	US-08-377-309-2	Sequence 2, Appli
36	1206.5	33.8	590	2	US-09-186-723-2	Sequence 2, Appli
37	1206.5	33.8	590	2	US-08-505-012-5	Sequence 5, Appli
38	1206.5	33.8	590	2	US-09-186-949A-3	Sequence 3, Appli
39	1206.5	33.8	590	2	US-08-758-757-2	Sequence 2, Appli
40	1206.5	33.8	590	2	US-09-187-978-2	Sequence 2, Appli
41	1206.5	33.8	590	2	US-10-115-701A-2	Sequence 2, Appli
42	1206.5	33.8	590	2	US-09-940-308A-2	Sequence 2, Appli
43	1206.5	33.8	590	2	US-09-940-308A-2	Sequence 2, Appli
44	1206.5	33.8	590	4	PCT-US96-00996-5	Sequence 5, Appli
45	1164.5	32.6	579	1	US-08-448-196A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne
; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: ST92007-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-256-938-4

Query Match 87.2%; Score 3112.5; DB 1; Length 787;
Best Local Similarity 94.6%; Pred. No. 2.3e-279;
Matches 596; Conservative 4; Mismatches 29; Indels 1; Gaps 1;

Qy 46 QAAKEFIAWLVKGRHGEGTFTSDVSSYLE-GQAAKEFIAWLVKGRDAHKSEVAHRFKDLG 104

```
Db        158 QGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQPGGGGDAHKSEVAHRFKDLG 217

Qy        105 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 277

Qy        165 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 337

Qy        225 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 397

Qy        285 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 457

Qy        345 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 517

Qy        405 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 577

Qy        465 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 637

Qy        525 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 697

Qy        585 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 757

Qy        645 EKCCKADDKETCFAEEGKKLVAASQAALGL 674
              ||||||||||||||||||||||||||||||
Db        758 EKCCKADDKETCFAEEGKKLVAASQAALGL 787


RESULT 2
US-08-797-689-16
; Sequence 16, Application US/08797689
; Patent No. 5876969
;  GENERAL INFORMATION:
;    APPLICANT:  Fleer, Reinhard
;    APPLICANT:  Fournier, Alain
;    APPLICANT:  Guitton, Jean-Dominique
;    APPLICANT:  Jung, Gerard
;    APPLICANT:  Yeh, Patrice
;    TITLE OF INVENTION:  NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;    TITLE OF INVENTION:  PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;    TITLE OF INVENTION:  CONTAINING SAID POLYPEPTIDES
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.1
;      SOFTWARE:  Word 5.1 (PatentIn)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/797,689
;      FILING DATE:  31-JAN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/256,927
;      FILING DATE:  28-JUL-1994
;      APPLICATION NUMBER:  FR 92/01064
;      FILING DATE:  31-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00085
;      FILING DATE:  28-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith Ph.D., Julie K.
;      REGISTRATION NUMBER:  P-38,619
;      REFERENCE/DOCKET NUMBER:  ST92006-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3839
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  787 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-797-689-16

  Query Match             87.2%;  Score 3112.5;  DB 1;  Length 787;
  Best Local Similarity   94.6%;  Pred. No. 2.3e-279;
  Matches  596;  Conservative    4;  Mismatches   29;  Indels    1;  Gaps    1;

Qy          46 QAAKEFIAWLVKGRHGEGTFTSDVSSYLE-GQAAKEFIAWLVKGRDAHKSEVAHRFKDLG 104
               | |    |   : | |     | : |:||       :|    | ||||||||||||||||
Db        158 QGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQPGGGGDAHKSEVAHRFKDLG 217

Qy        105 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 277

Qy        165 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 337

Qy        225 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 397

Qy        285 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 457

Qy        345 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 517

Qy        405 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 577

Qy        465 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 637

Qy        525 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 697

Qy        585 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 757

Qy        645 EKCCKADDKETCFAEEGKKLVAASQAALGL 674
              ||||||||||||||||||||||||||||||
Db        758 EKCCKADDKETCFAEEGKKLVAASQAALGL 787


RESULT 3
```

```
US-09-984-186-16
; Sequence 16, Application US/09984186
; Patent No. 6686179
;  GENERAL INFORMATION:
;       APPLICANT: Pleer, Reinhard
;                  Fournier, Alain
;                  Guitton, Jean-Dominique
;                  Jung, Gerard
;                  Yeh, Patrice
;       TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                           PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                           CONTAINING SAID POLYPEPTIDES
;       NUMBER OF SEQUENCES: 36
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Rhone-Poulenc Rorer Inc.
;            STREET: 500 Arcola Road, 3C43
;            CITY: Collegeville
;            STATE: PA
;            COUNTRY: USA
;            ZIP: 19426
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: Macintosh
;            OPERATING SYSTEM: System 7.1
;            SOFTWARE: Word 5.1 (PatentIn)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/984,186
;            FILING DATE: 29-Oct-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/797,689
;            FILING DATE: 31-JAN-1997
;            APPLICATION NUMBER: US 08/256,927
;            FILING DATE: 28-JUL-1994
;            APPLICATION NUMBER: FR 92/01064
;            FILING DATE: 31-JAN-1992
;            APPLICATION NUMBER: PCT/FR93/00085
;            FILING DATE: 28-JAN-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Smith Ph.D., Julie K.
;            REGISTRATION NUMBER: P-38,619
;            REFERENCE/DOCKET NUMBER: ST92006-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (610) 454-3839
;            TELEFAX: (610) 454-3808
;  INFORMATION FOR SEQ ID NO: 16:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 787 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-984-186-16

  Query Match             87.2%;  Score 3112.5;  DB 2;  Length 787;
  Best Local Similarity   94.6%;  Pred. No. 2.3e-279;
  Matches  596;  Conservative    4;  Mismatches   29;  Indels    1;  Gaps    1;

Qy          46 QAAKEFIAWLVKGRHGEGTFTSDVSSYLE-GQAAKEFIAWLVKGRDAHKSEVAHRFKDLG 104
               |||     |   : | |     | : |:||       :|    | ||||||||||||||
Db         158 QGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQPGGGGDAHKSEVAHRFKDLG 217

Qy         105 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         218 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 277

Qy         165 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 224
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         278 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 337

Qy         225 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 284
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         338 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 397

Qy         285 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 344
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         398 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 457

Qy         345 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 404
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         458 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 517

Qy         405 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 464
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         518 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 577

Qy         465 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 524
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         578 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 637

Qy         525 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 584
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         638 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 697

Qy         585 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 644
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         698 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 757

Qy         645 EKCCKADDKETCFAEEGKKLVAASQAALGL 674
               |||||||||||||||||||||||||||||
Db         758 EKCCKADDKETCFAEEGKKLVAASQAALGL 787


RESULT 4
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 977
;   LENGTH: 609
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

  Query Match             87.1%;  Score 3108;  DB 2;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-279;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy         149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy         209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy         269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263
```

```
Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

RESULT 5
US-09-919-039-370
; Sequence 370, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6727066 088957CD1
US-09-919-039-370

Query Match 87.1%; Score 3108; DB 2; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.1e-279;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

RESULT 6
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-2

```
  Query Match             87.1%;  Score 3108;  DB 1;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-279;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 7
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
;   GENERAL INFORMATION:
;        APPLICANT: Fleer, Reinhard
;                   Fournier, Alain
;                   Guitton, Jean-Dominique
;                   Jung, Gerard
;                   Yeh, Patrice
;        TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                            PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                            CONTAINING SAID POLYPEPTIDES
;        NUMBER OF SEQUENCES: 36
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rhone-Poulenc Rorer Inc.
;             STREET: 500 Arcola Road, 3C43
;             CITY: Collegeville
;             STATE: PA
;             COUNTRY: USA
;             ZIP: 19426
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: Macintosh
;             OPERATING SYSTEM: System 7.1
;             SOFTWARE: Word 5.1 (PatentIn)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/984,186
;             FILING DATE: 29-Oct-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/797,689
;             FILING DATE: 31-JAN-1997
;             APPLICATION NUMBER: US 08/256,927
;             FILING DATE: 28-JUL-1994
;             APPLICATION NUMBER: FR 92/01064
;             FILING DATE: 31-JAN-1992
;             APPLICATION NUMBER: PCT/FR93/00085
;             FILING DATE: 28-JAN-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Smith Ph.D., Julie K.
;             REGISTRATION NUMBER: P-38,619
;             REFERENCE/DOCKET NUMBER: ST92006-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (610) 454-3839
;             TELEFAX: (610) 454-3808
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 610 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

  Query Match             87.1%;  Score 3108;  DB 2;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-279;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

```
RESULT 8
US-09-949-016-11170
; Sequence 11170, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11170
;   LENGTH: 622
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-11170

  Query Match             87.1%;  Score 3108;  DB 2;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-279;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 96

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 156

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 216

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 276

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 336

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 396

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 456

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 516

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 576

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        577 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 622


RESULT 9
US-08-256-938-2
; Sequence 2, Application US/08256938
; Patent No. 5665863
;  GENERAL INFORMATION:
;    APPLICANT:  Yeh, Patrice
;    TITLE OF INVENTION:  NEW POLYPEPTIDES HAVING GRANULOCYTE
;    TITLE OF INVENTION:  COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
;    TITLE OF INVENTION:  PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.1
;      SOFTWARE:  Word 5.0 (PatentIn)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/256,938
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 92/01065
;      FILING DATE:  31-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goodman, Rosanne
;      REGISTRATION NUMBER:  32,534
;      REFERENCE/DOCKET NUMBER:  ST92007-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3817
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  783 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-256-938-2

  Query Match             87.1%;  Score 3108;  DB 1;  Length 783;
  Best Local Similarity   100.0%;  Pred. No. 6e-279;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443
```

```
Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 10
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
;  GENERAL INFORMATION:
;    APPLICANT:  Lichenstein, Henri
;    APPLICANT:  Lyons, David
;    APPLICANT:  Wurfel, Mark
;    APPLICANT:  Wright, Samuel
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/222,619
;      FILING DATE:
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  609 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-222-619-3

  Query Match             87.0%;  Score 3104;  DB 1;  Length 609;
  Best Local Similarity   99.8%;  Pred. No. 9.6e-279;
  Matches  585;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFKQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 11
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
;  GENERAL INFORMATION:
;    APPLICANT:  Sreekrishna, Kotikanyadan
;    APPLICANT:  Barr, Kathryn A.
;    APPLICANT:  Brierley, Russell A.
;    APPLICANT:  Thill, Gregory P.
;    APPLICANT:  Tschopp, Juerg F.
;    TITLE OF INVENTION:  EXPRESSION OF HUMAN SERUM ALBUMIN IN
;    TITLE OF INVENTION:  PICHIA PASTORIS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/433,037
;      FILING DATE:  03-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9108Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  609 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-433-037-4

  Query Match             87.0%;  Score 3104;  DB 1;  Length 609;
  Best Local Similarity   99.8%;  Pred. No. 9.6e-279;
  Matches  585;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKVSSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 12
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
;  APPLICANT:  Mary Ellen Digan
;  APPLICANT:  Philip Lake
;  APPLICANT:  Hermann Gram
;  TITLE OF INVENTION:  Fusion Polypeptides
;  FILE REFERENCE:  600-7244/CPA
;  CURRENT APPLICATION NUMBER: US/08/897,956A
;  CURRENT FILING DATE:  1997-07-21
;  PRIOR APPLICATION NUMBER:  60/022,689
;  PRIOR FILING DATE:  1996-07-26
;  NUMBER OF SEQ ID NOS:  38
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 609
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-08-897-956A-2

  Query Match             87.0%;  Score 3104;  DB 2;  Length 609;
  Best Local Similarity   99.8%;  Pred. No. 9.6e-279;
  Matches  585;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFKQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 13
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
;  GENERAL INFORMATION:
;    APPLICANT:  AMGEN INC.
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04075
;      FILING DATE:
;      CLASSIFICATION:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  609 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
PCT-US95-04075-3

  Query Match             87.0%;  Score 3104;  DB 4;  Length 609;
  Best Local Similarity   99.8%;  Pred. No. 9.6e-279;
  Matches  585;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83
```

```
Qy        149 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        209 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        269 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        329 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        389 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        449 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 508
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFKQLGEYKFQNALLVRYTKKVPQVS 443

Qy        509 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        569 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        629 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

RESULT 14
US-08-897-956A-3
; Sequence 3, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 978
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion polypeptide
US-08-897-956A-3

Query Match 87.0%; Score 3103.5; DB 2; Length 978;
Best Local Similarity 98.8%; Pred. No. 2.2e-278;
Matches 586; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

```
Qy         84 WLVK---GRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAK 140
              ||   | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 WLASGGGGSDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAK 262

Qy        141 TCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPN 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 TCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPN 322

Qy        201 LPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQA 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 LPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQA 382

Qy        261 ADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEV 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 ADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEV 442

Qy        321 SKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 SKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA 502

Qy        381 EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAK 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAK 562

Qy        441 TYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRY 500
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        563 TYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFKQLGEYKFQNALLVRY 622

Qy        501 TKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSD 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 TKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSD 682

Qy        561 RVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVE 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 RVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVE 742

Qy        621 LVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALG 673
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 LVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALG 795
```

RESULT 15
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain

```
;      REGISTRATION NUMBER:  24864
;      REFERENCE/DOCKET NUMBER:  92H832
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 665 2400
;      TELEFAX:  (908) 771 6159
;      TELEX:  219484
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  369..419
;      OTHER INFORMATION:  /note= "Alternative C-termini of
;      OTHER INFORMATION:  HSA(1-n)"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..585
;      OTHER INFORMATION:  /note= "Amino acid sequence of
;      OTHER INFORMATION:  natural HSA"
US-08-153-799-14

  Query Match             87.0%;  Score 3103;  DB 1;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-278;
  Matches  585;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          90 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 149
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 60

Qy         150 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 209
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 120

Qy         210 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 269
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 180

Qy         270 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 329
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240

Qy         330 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 389
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300

Qy         390 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 449
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 360

Qy         450 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 509
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 420

Qy         510 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 569
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Qy         570 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 629
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540

Qy         630 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674
               |||||||||||||||||||||||||||||||||||||||||||||
Db         541 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 585


Search completed: April 19, 2006, 12:11:34
```

Job time : 44.6037 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:29:13 ; Search time 142.172 Seconds
(without alignments)
1980.821 Million cell updates/sec

Title: US-10-775-180-447
Perfect score: 3568
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....TCFAEEGKKLVAASQAALGL 674

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3568	100.0	674	5	US-10-775-180-447	Sequence 447, App
2	3568	100.0	674	5	US-10-775-204-1280	Sequence 1280, Ap
3	3444.5	96.5	669	5	US-10-775-180-419	Sequence 419, App
4	3444.5	96.5	669	5	US-10-775-204-1231	Sequence 1231, Ap
5	3444	96.5	730	5	US-10-775-180-610	Sequence 610, App
6	3444	96.5	730	5	US-10-775-204-1622	Sequence 1622, Ap
7	3438.5	96.4	669	5	US-10-775-180-425	Sequence 425, App
8	3438.5	96.4	669	5	US-10-775-204-1237	Sequence 1237, Ap
9	3438	96.4	730	5	US-10-775-180-612	Sequence 612, App
10	3438	96.4	730	5	US-10-775-204-1624	Sequence 1624, Ap
11	3432.5	96.2	669	5	US-10-775-180-420	Sequence 420, App
12	3432.5	96.2	669	5	US-10-775-180-421	Sequence 421, App
13	3432.5	96.2	669	5	US-10-775-180-423	Sequence 423, App
14	3432.5	96.2	669	5	US-10-775-180-424	Sequence 424, App
15	3432.5	96.2	669	5	US-10-775-204-1232	Sequence 1232, Ap
16	3432.5	96.2	669	5	US-10-775-204-1233	Sequence 1233, Ap
17	3432.5	96.2	669	5	US-10-775-204-1235	Sequence 1235, Ap
18	3432.5	96.2	669	5	US-10-775-204-1236	Sequence 1236, Ap
19	3427	96.0	668	5	US-10-775-180-609	Sequence 609, App
20	3427	96.0	668	5	US-10-775-204-1621	Sequence 1621, Ap
21	3422	95.9	662	5	US-10-775-180-611	Sequence 611, App
22	3422	95.9	662	5	US-10-775-204-1623	Sequence 1623, Ap
23	3421	95.9	668	5	US-10-775-180-613	Sequence 613, App
24	3421	95.9	668	5	US-10-775-204-1625	Sequence 1625, Ap
25	3420.5	95.9	664	5	US-10-775-180-598	Sequence 598, App
26	3420.5	95.9	664	5	US-10-775-204-1607	Sequence 1607, Ap
27	3418.5	95.8	663	5	US-10-775-180-600	Sequence 600, App
28	3418.5	95.8	663	5	US-10-775-204-1609	Sequence 1609, Ap
29	3416	95.7	662	5	US-10-775-180-614	Sequence 614, App
30	3416	95.7	662	5	US-10-775-204-1626	Sequence 1626, Ap
31	3414.5	95.7	664	5	US-10-775-180-599	Sequence 599, App
32	3414.5	95.7	664	5	US-10-775-204-1608	Sequence 1608, Ap
33	3413.5	95.7	667	5	US-10-775-180-422	Sequence 422, App
34	3413.5	95.7	667	5	US-10-775-204-1234	Sequence 1234, Ap
35	3412.5	95.6	663	5	US-10-775-180-601	Sequence 601, App
36	3412.5	95.6	663	5	US-10-775-204-1610	Sequence 1610, Ap
37	3278	91.9	654	5	US-10-775-180-574	Sequence 574, App
38	3278	91.9	654	5	US-10-775-204-1559	Sequence 1559, Ap
39	3275.5	91.8	655	5	US-10-775-180-623	Sequence 623, App
40	3275.5	91.8	655	5	US-10-775-204-1640	Sequence 1640, Ap
41	3267	91.6	682	5	US-10-775-204-1655	Sequence 1655, Ap
42	3265	91.5	639	5	US-10-775-180-131	Sequence 131, App
43	3265	91.5	639	5	US-10-775-204-417	Sequence 417, App
44	3261.5	91.4	700	5	US-10-775-204-1620	Sequence 1620, Ap
45	3259	91.3	639	5	US-10-775-180-129	Sequence 129, App

ALIGNMENTS

RESULT 1
US-10-775-180-447
; Sequence 447, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 447
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-180-447

Query Match 100.0%; Score 3568; DB 5; Length 674;
Best Local Similarity 100.0%; Pred. No. 1.4e-270;
Matches 674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        661 GKKLVAASQAALGL 674


RESULT 2
US-10-775-204-1280
; Sequence 1280, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1280
;   LENGTH: 674
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1280

  Query Match            100.0%;  Score 3568;  DB 5;  Length 674;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-270;
  Matches  674;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        661 GKKLVAASQAALGL 674


RESULT 3
US-10-775-180-419
; Sequence 419, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
```

```
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 419
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-419

  Query Match             96.5%;  Score 3444.5;  DB 5;  Length 669;
  Best Local Similarity   97.3%;  Pred. No. 6.8e-261;
  Matches  654;  Conservative    4;  Mismatches    5;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||||||||||||||||||||||||||||||||||
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669


RESULT 4
US-10-775-204-1231
; Sequence 1231, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1231
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1231

  Query Match             96.5%;  Score 3444.5;  DB 5;  Length 669;
  Best Local Similarity   97.3%;  Pred. No. 6.8e-261;
  Matches  654;  Conservative    4;  Mismatches    5;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||||||||||||||||||||||||||||||||||
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237
```

```
Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669
```

RESULT 5
US-10-775-180-610
; Sequence 610, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 610
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-180-610

```
  Query Match           96.5%;  Score 3444;  DB 5;  Length 730;
  Best Local Similarity  97.8%;  Pred. No. 8.3e-261;
  Matches  652;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 183

Qy        128 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 243

Qy        188 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 303

Qy        248 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 363

Qy        308 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 423

Qy        368 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 483

Qy        428 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 543

Qy        488 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 603

Qy        548 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 663

Qy        608 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 723

Qy        668 SQAALGL 674
              |||||||
Db        724 SQAALGL 730
```

RESULT 6
US-10-775-204-1622
; Sequence 1622, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10

```
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1622
;   LENGTH: 730
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1622

  Query Match             96.5%;  Score 3444;  DB 5;  Length 730;
  Best Local Similarity   97.8%;  Pred. No. 8.3e-261;
  Matches  652;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :      ||||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 183

Qy        128 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 243

Qy        188 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 303

Qy        248 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 363

Qy        308 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 423

Qy        368 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 483

Qy        428 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 543

Qy        488 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 603

Qy        548 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 663

Qy        608 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 723

Qy        668 SQAALGL 674
              |||||||
Db        724 SQAALGL 730


RESULT 7
US-10-775-180-425
; Sequence 425, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 425
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-425

  Query Match             96.4%;  Score 3438.5;  DB 5;  Length 669;
  Best Local Similarity   97.2%;  Pred. No. 2e-260;
  Matches  653;  Conservative    4;  Mismatches    6;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :          |||||||||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669


RESULT 8
US-10-775-204-1237
; Sequence 1237, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1237
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1237

  Query Match             96.4%;  Score 3438.5;  DB 5;  Length 669;
  Best Local Similarity   97.2%;  Pred. No. 2e-260;
  Matches  653;  Conservative    4;  Mismatches    6;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||||||||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669


RESULT 9
US-10-775-180-612
; Sequence 612, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
```

; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 612
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-180-612

Query Match 96.4%; Score 3438; DB 5; Length 730;
Best Local Similarity 97.6%; Pred. No. 2.5e-260;
Matches 651; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

```
Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||| ||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 183

Qy        128 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 243

Qy        188 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 303

Qy        248 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 363

Qy        308 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 423

Qy        368 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 483

Qy        428 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 543

Qy        488 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 603

Qy        548 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 663

Qy        608 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 723

Qy        668 SQAALGL 674
              |||||||
Db        724 SQAALGL 730
```

RESULT 10
US-10-775-204-1624
; Sequence 1624, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1624
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-1624

Query Match 96.4%; Score 3438; DB 5; Length 730;
Best Local Similarity 97.6%; Pred. No. 2.5e-260;
Matches 651; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

```
Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||| ||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED 183

Qy        128 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER 243

Qy        188 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFA 303

Qy        248 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 KRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR 363

Qy        308 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKEC 423

Qy        368 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH 483

Qy        428 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQL 543

Qy        488 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLN 603
```

```
Qy        548 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 QLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSE 663

Qy        608 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 KERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAA 723

Qy        668 SQAALGL 674
              |||||||
Db        724 SQAALGL 730


RESULT 11
US-10-775-180-420
; Sequence 420, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 420
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-420

  Query Match             96.2%;  Score 3432.5;  DB 5;  Length 669;
  Best Local Similarity   97.0%;  Pred. No. 5.9e-260;
  Matches  652;  Conservative    4;  Mismatches    7;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         |||||| |||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669


RESULT 12
US-10-775-180-421
; Sequence 421, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 421
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-421
```

```
  Query Match             96.2%;  Score 3432.5;  DB 5;  Length 669;
  Best Local Similarity   97.0%;  Pred. No. 5.9e-260;
  Matches  652;  Conservative    4;  Mismatches    7;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||| |||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669


RESULT 13
US-10-775-180-423
; Sequence 423, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 423
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-423

  Query Match             96.2%;  Score 3432.5;  DB 5;  Length 669;
  Best Local Similarity   97.0%;  Pred. No. 5.9e-260;
  Matches  652;  Conservative    4;  Mismatches    7;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||| |||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHSEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669
```

```
RESULT 14
US-10-775-180-424
; Sequence 424, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 424
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-424

  Query Match            96.2%;  Score 3432.5;  DB 5;  Length 669;
  Best Local Similarity  97.0%;  Pred. No. 5.9e-260;
  Matches  652;  Conservative    4;  Mismatches    7;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||| |||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669


RESULT 15
US-10-775-204-1232
; Sequence 1232, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1232
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1232

  Query Match            96.2%;  Score 3432.5;  DB 5;  Length 669;
  Best Local Similarity  97.0%;  Pred. No. 5.9e-260;
  Matches  652;  Conservative    4;  Mismatches    7;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||| |||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQ 117

Qy        123 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK 177

Qy        183 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPE 237

Qy        243 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 LLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA 297

Qy        303 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISS 357

Qy        363 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 KLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE 417

Qy        423 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE 477

Qy        483 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYL 537

Qy        543 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 SVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI 597

Qy        603 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGK 657

Qy        663 KLVAASQAALGL 674
              ||||||||||||
Db        658 KLVAASQAALGL 669
```

Search completed: April 19, 2006, 12:35:48
Job time : 144.172 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:31:03 ; Search time 22.4986 Seconds
(without alignments)
1318.215 Million cell updates/sec

Title: US-10-775-180-447
Perfect score: 3568
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....TCFAEEGKKLVAASQAALGL 674

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225428 seqs, 44002918 residues

Total number of hits satisfying chosen parameters: 225428

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /SIDS5/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /SIDS5/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /SIDS5/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /SIDS5/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /SIDS5/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /SIDS5/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /SIDS5/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /SIDS5/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3568	100.0	674	7	US-11-175-690-206	Sequence 206, App
2	3568	100.0	915	7	US-11-175-690-208	Sequence 208, App
3	3401	95.3	658	7	US-11-175-690-210	Sequence 210, App
4	3395	95.2	654	7	US-11-175-690-219	Sequence 219, App
5	3395	95.2	656	7	US-11-175-690-225	Sequence 225, App
6	3393	95.1	650	7	US-11-175-690-209	Sequence 209, App
7	3392.5	95.1	655	7	US-11-175-690-220	Sequence 220, App
8	3392.5	95.1	657	7	US-11-175-690-216	Sequence 216, App
9	3392.5	95.1	657	7	US-11-175-690-303	Sequence 303, App
10	3391.5	95.1	653	7	US-11-175-690-215	Sequence 215, App
11	3391	95.0	652	7	US-11-175-690-218	Sequence 218, App
12	3390.5	95.0	659	7	US-11-175-690-221	Sequence 221, App
13	3389	95.0	648	7	US-11-175-690-214	Sequence 214, App
14	3387.5	94.9	651	7	US-11-175-690-224	Sequence 224, App
15	3386.5	94.9	647	7	US-11-175-690-212	Sequence 212, App
16	3386.5	94.9	649	7	US-11-175-690-213	Sequence 213, App
17	3386	94.9	646	7	US-11-175-690-223	Sequence 223, App
18	3214	90.1	779	7	US-11-175-690-205	Sequence 205, App
19	3198	89.6	646	7	US-11-175-690-276	Sequence 276, App
20	3190	89.4	678	7	US-11-175-690-274	Sequence 274, App
21	3182	89.2	642	7	US-11-175-690-238	Sequence 238, App
22	3176	89.0	642	7	US-11-175-690-233	Sequence 233, App
23	3171.5	88.9	647	7	US-11-175-690-242	Sequence 242, App
24	3159	88.5	636	7	US-11-175-690-268	Sequence 268, App
25	3158	88.5	636	7	US-11-175-690-278	Sequence 278, App
26	3154	88.4	636	7	US-11-175-690-240	Sequence 240, App
27	3144.5	88.1	693	7	US-11-175-690-199	Sequence 199, App
28	3137.5	87.9	688	7	US-11-175-690-198	Sequence 198, App
29	3134.5	87.9	637	7	US-11-175-690-266	Sequence 266, App
30	3131.5	87.8	629	7	US-11-175-690-562	Sequence 562, App
31	3126	87.6	728	7	US-11-175-690-244	Sequence 244, App
32	3126	87.6	728	7	US-11-175-690-246	Sequence 246, App
33	3126	87.6	728	7	US-11-175-690-248	Sequence 248, App
34	3125.5	87.6	672	7	US-11-175-690-200	Sequence 200, App
35	3122.5	87.5	673	7	US-11-175-690-201	Sequence 201, App
36	3118.5	87.4	667	7	US-11-175-690-227	Sequence 227, App
37	3116.5	87.3	673	7	US-11-175-690-217	Sequence 217, App
38	3116.5	87.3	673	7	US-11-175-690-231	Sequence 231, App
39	3116	87.3	728	7	US-11-175-690-254	Sequence 254, App
40	3115.5	87.3	663	7	US-11-175-690-284	Sequence 284, App
41	3112.5	87.2	634	7	US-11-175-690-207	Sequence 207, App
42	3112.5	87.2	665	7	US-11-175-690-282	Sequence 282, App
43	3111.5	87.2	661	7	US-11-175-690-281	Sequence 281, App
44	3110	87.2	742	7	US-11-175-690-525	Sequence 525, App
45	3109.5	87.1	638	7	US-11-175-690-229	Sequence 229, App

ALIGNMENTS

RESULT 1
US-11-175-690-206
; Sequence 206, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 206
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-206

Query Match 100.0%; Score 3568; DB 7; Length 674;
Best Local Similarity 100.0%; Pred. No. 2.3e-274;
Matches 674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        661 GKKLVAASQAALGL 674
```

RESULT 2
US-11-175-690-208
; Sequence 208, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 208
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-208

```
  Query Match           100.0%;  Score 3568;  DB 7;  Length 915;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-274;
  Matches  674;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        661 GKKLVAASQAALGL 674
```

RESULT 3
US-11-175-690-210
; Sequence 210, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23

```
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 210
;   LENGTH: 658
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-210

  Query Match             95.3%;  Score 3401;  DB 7;  Length 658;
  Best Local Similarity   96.1%;  Pred. No. 3.6e-261;
  Matches  648;  Conservative    1;  Mismatches    9;  Indels   16;  Gaps    1;

Qy          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                           | |  :|   |  ||||||||||||||||||||||||||||||
Db         60 --------------DAHKSEVAHRFKDLDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 104

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 164

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 224

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 284

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 344

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 404

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 464

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 524

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 584

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        585 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 644

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        645 GKKLVAASQAALGL 658


RESULT 4
US-11-175-690-219
; Sequence 219, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 219
;   LENGTH: 654
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-219

  Query Match             95.2%;  Score 3395;  DB 7;  Length 654;
  Best Local Similarity   96.1%;  Pred. No. 1.1e-260;
  Matches  648;  Conservative    1;  Mismatches    5;  Indels   20;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                           | | :|   |  ||||||||||||||||||||||||||||||
Db         60 --------------DAHKSEVAHRFKDLDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 100

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 160

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 220

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 280

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 340

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 400

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 460

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 520

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 580

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 640

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        641 GKKLVAASQAALGL 654


RESULT 5
US-11-175-690-225
; Sequence 225, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 225
;   LENGTH: 656
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-225

  Query Match             95.2%;  Score 3395;  DB 7;  Length 656;
  Best Local Similarity   95.8%;  Pred. No. 1.1e-260;
  Matches  646;  Conservative    3;  Mismatches    7;  Indels   18;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                        : :  |  |       :||||||||||||||||||||||||||||||
Db         61 AH-----------KSEVAHRF-------KDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 102

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 162

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 222

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 282

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 342

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 402

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 462

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 522

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 582

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 642

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        643 GKKLVAASQAALGL 656


RESULT 6
US-11-175-690-209
; Sequence 209, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 209
;   LENGTH: 650
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-209

  Query Match             95.1%;  Score 3393;  DB 7;  Length 650;
  Best Local Similarity   95.7%;  Pred. No. 1.5e-260;
  Matches  645;  Conservative    0;  Mismatches    5;  Indels   24;  Gaps    1;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                                     |  ||||||||||||||||||||||||||||||
Db         61 AH-----------------------KSEDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 96

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 156

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
```

```
Db        157 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 216

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300

Db        217 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 276

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Db        277 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 336

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Db        337 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 396

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Db        397 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 456

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Db        457 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 516

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Db        517 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 576

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Db        577 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 636

Qy        661 GKKLVAASQAALGL 674

Db        637 GKKLVAASQAALGL 650
```

RESULT 7
US-11-175-690-220
; Sequence 220, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 220
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-220

Query Match 95.1%; Score 3392.5; DB 7; Length 655;
Best Local Similarity 95.8%; Pred. No. 1.7e-260;
Matches 646; Conservative 2; Mismatches 7; Indels 19; Gaps 2;

```
Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120

Db         61 AH-----------KSEVAHRF--------DAHKSEVAHRFKDLGEENFKALVLIAFAQYL 101

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180

Db        102 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 161

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240

Db        162 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 221

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300

Db        222 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 281

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360

Db        282 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 341

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420

Db        342 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 401

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480

Db        402 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 461

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540

Db        462 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 521

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600

Db        522 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 581

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660

Db        582 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 641

Qy        661 GKKLVAASQAALGL 674

Db        642 GKKLVAASQAALGL 655
```

RESULT 8
US-11-175-690-216
; Sequence 216, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172

```
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 216
;   LENGTH: 657
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-216

  Query Match             95.1%;  Score 3392.5;  DB 7;  Length 657;
  Best Local Similarity   96.1%;  Pred. No. 1.7e-260;
  Matches  648;  Conservative    1;  Mismatches    8;  Indels   17;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                          | |  :|    |  ||||||||||||||||||||||||||||||||
Db         60 ---------------DAHKSEVAHRFKD-DAHKSEVAHRFKDLGEENFKALVLIAFAQYL 103

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 163

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 223

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 283

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 343

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 403

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 463

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 523

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 583

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 643

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        644 GKKLVAASQAALGL 657


RESULT 9
US-11-175-690-303
; Sequence 303, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 303
;   LENGTH: 657
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-303

  Query Match             95.1%;  Score 3392.5;  DB 7;  Length 657;
  Best Local Similarity   96.1%;  Pred. No. 1.7e-260;
  Matches  648;  Conservative    1;  Mismatches    8;  Indels   17;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                          | |  :|    |  ||||||||||||||||||||||||||||||||
Db         60 ---------------DAHKSEVAHRFKD-DAHKSEVAHRFKDLGEENFKALVLIAFAQYL 103

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 163

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 223

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 283

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 343

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 403

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 463

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 523

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 583

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 643
```

```
Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        644 GKKLVAASQAALGL 657


RESULT 10
US-11-175-690-215
; Sequence 215, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 215
;   LENGTH: 653
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-215

  Query Match             95.1%;  Score 3391.5;  DB 7;  Length 653;
  Best Local Similarity   96.0%;  Pred. No. 2e-260;
  Matches  647;  Conservative    1;  Mismatches    5;  Indels   21;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                         | |  :|      ||||||||||||||||||||||||||||||||
Db         60 --------------DAHKSEVA-----HDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 99

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 159

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 219

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 279

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 339

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 399

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 459

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 519

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 579

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 639

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        640 GKKLVAASQAALGL 653


RESULT 11
US-11-175-690-218
; Sequence 218, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 218
;   LENGTH: 652
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-218

  Query Match             95.0%;  Score 3391;  DB 7;  Length 652;
  Best Local Similarity   96.0%;  Pred. No. 2.2e-260;
  Matches  647;  Conservative    1;  Mismatches    4;  Indels   22;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                         | |  :|      ||||||||||||||||||||||||||||||||
Db         60 --------------DAHKSEVA------DAHKSEVAHRFKDLGEENFKALVLIAFAQYL 98

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 158

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 218
```

```
Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 278

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 338

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 398

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 458

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 518

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 578

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 638

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        639 GKKLVAASQAALGL 652


RESULT 12
US-11-175-690-221
; Sequence 221, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 221
;   LENGTH: 659
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-221

  Query Match             95.0%;  Score 3390.5;  DB 7;  Length 659;
  Best Local Similarity   96.0%;  Pred. No. 2.5e-260;
  Matches  648;  Conservative    1;  Mismatches    9;  Indels   17;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKG-RDAHKSEVAHRFKDLGEENFKALVLIAFAQY 119
                           ||  :|    |   ||||||||||||||||||||||||||||||
Db         60 --------------DAHKSEVAHRFKDLGDAHKSEVAHRFKDLGEENFKALVLIAFAQY 104

Qy        120 LQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADC 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 LQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADC 164

Qy        180 CAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFY 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 CAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFY 224

Qy        240 APELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERA 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 APELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERA 284

Qy        300 FKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 FKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDS 344

Qy        360 ISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMF 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 ISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMF 404

Qy        420 LYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQ 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 LYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQ 464

Qy        480 NCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 NCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAE 524

Qy        540 DYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFH 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 DYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFH 584

Qy        600 ADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAE 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        585 ADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAE 644

Qy        660 EGKKLVAASQAALGL 674
              ||||||||||||||
Db        645 EGKKLVAASQAALGL 659


RESULT 13
US-11-175-690-214
; Sequence 214, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
```

```
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 214
;   LENGTH: 648
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-214

  Query Match             95.0%;  Score 3389;  DB 7;  Length 648;
  Best Local Similarity   95.5%;  Pred. No. 3.2e-260;
  Matches  644;  Conservative    1;  Mismatches    3;  Indels   26;  Gaps    1;

Qy          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                                      :|||||||||||||||||||||||||||||||||
Db         61 AH-------------------------KDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 94

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 154

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 214

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 274

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 334

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 394

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 454

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 514

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 574

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 634

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        635 GKKLVAASQAALGL 648


RESULT 14
US-11-175-690-224
; Sequence 224, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 224
;   LENGTH: 651
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-224

  Query Match             94.9%;  Score 3387.5;  DB 7;  Length 651;
  Best Local Similarity   95.8%;  Pred. No. 4.2e-260;
  Matches  646;  Conservative    1;  Mismatches    4;  Indels   23;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                |:|                    ||||||||||||||||||||||||||||||||||
Db         61 AH---KSEV-------------------DAHKSEVAHRFKDLGEENFKALVLIAFAQYL 97

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 157

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 217

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 277

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 337

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 397

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 457

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 517

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 577

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 637

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        638 GKKLVAASQAALGL 651
```

```
RESULT 15
US-11-175-690-212
; Sequence 212, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 212
;   LENGTH: 647
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-212

  Query Match             94.9%;  Score 3386.5;  DB 7;  Length 647;
  Best Local Similarity   95.5%;  Pred. No. 5e-260;
  Matches  644;  Conservative    0;  Mismatches    3;  Indels   27;  Gaps    1;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 120
                                           |||||||||||||||||||||||||||||
Db         60 -------------------------DAHDAHKSEVAHRFKDLGEENFKALVLIAFAQYL 93

Qy        121 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 QQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCC 153

Qy        181 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 AKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYA 213

Qy        241 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 PELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAF 273

Qy        301 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 KAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSI 333

Qy        361 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 SSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFL 393

Qy        421 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 YEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQN 453

Qy        481 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 CELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAED 513

Qy        541 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 YLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHA 573

Qy        601 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 DICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE 633

Qy        661 GKKLVAASQAALGL 674
              ||||||||||||||
Db        634 GKKLVAASQAALGL 647


Search completed: April 19, 2006, 12:36:41
Job time : 24.4986 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 11:56:31 ; Search time 153.004 Seconds
(without alignments)
1852.232 Million cell updates/sec

Title: US-10-775-180-447_COPY_30_674
Perfect score: 3417
Sequence: 1 HGEGTFTSDVSSYLEGQAAK.....TCFAEEGKKLVAASQAALGL 645

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3417	100.0	662	7	ADF16526	Adf16526 Human alb
2	3417	100.0	662	7	ADH21814	Adh21814 Human alb
3	3417	100.0	663	7	ADF16512	Adf16512 Human alb
4	3417	100.0	663	7	ADH21803	Adh21803 Human alb
5	3417	100.0	664	7	ADF16510	Adf16510 Human alb
6	3417	100.0	664	7	ADH21801	Adh21801 Human alb
7	3417	100.0	668	7	ADF16524	Adf16524 Human alb
8	3417	100.0	668	7	ADH21812	Adh21812 Human alb
9	3417	100.0	669	7	ADF16144	Adf16144 Human alb
10	3417	100.0	669	7	ADH21622	Adh21622 Human alb
11	3417	100.0	674	7	ADF16193	Adf16193 Human alb
12	3417	100.0	674	7	ADH21650	Adh21650 Human alb
13	3417	100.0	674	9	ADW45202	Adw45202 K. lactis
14	3417	100.0	730	7	ADF16525	Adf16525 Human alb
15	3417	100.0	730	7	ADH21813	Adh21813 Human alb
16	3417	100.0	915	9	ADW45204	Adw45204 K. lactis
17	3411	99.8	662	7	ADF16529	Adf16529 Human alb
18	3411	99.8	662	7	ADH21817	Adh21817 Human alb
19	3411	99.8	663	7	ADF16513	Adf16513 Human alb
20	3411	99.8	663	7	ADH21804	Adh21804 Human alb
21	3411	99.8	664	7	ADF16511	Adf16511 Human alb
22	3411	99.8	664	7	ADH21802	Adh21802 Human alb
23	3411	99.8	668	7	ADF16528	Adf16528 Human alb
24	3411	99.8	668	7	ADH21816	Adh21816 Human alb
25	3411	99.8	669	7	ADF16150	Adf16150 Human alb
26	3411	99.8	669	7	ADH21628	Adh21628 Human alb
27	3411	99.8	730	7	ADF16527	Adf16527 Human alb
28	3411	99.8	730	7	ADH21815	Adh21815 Human alb
29	3405	99.6	669	7	ADF16149	Adf16149 Human alb
30	3405	99.6	669	7	ADF16148	Adf16148 Human alb
31	3405	99.6	669	7	ADF16145	Adf16145 Human alb
32	3405	99.6	669	7	ADF16146	Adf16146 Human alb
33	3405	99.6	669	7	ADH21624	Adh21624 Human alb
34	3405	99.6	669	7	ADH21626	Adh21626 Human alb
35	3405	99.6	669	7	ADH21623	Adh21623 Human alb
36	3405	99.6	669	7	ADH21627	Adh21627 Human alb
37	3397	99.4	667	7	ADF16147	Adf16147 Human alb
38	3397	99.4	667	7	ADH21625	Adh21625 Human alb
39	3265	95.6	639	7	ADF15119	Adf15119 Human alb
40	3265	95.6	639	7	ADH21334	Adh21334 Human alb
41	3259	95.4	639	7	ADF15116	Adf15116 Human alb
42	3259	95.4	639	7	ADH21332	Adh21332 Human alb
43	3259	95.4	700	7	ADF16523	Adf16523 Human alb
44	3254	95.2	646	9	ADW45219	Adw45219 K. lactis
45	3253.5	95.2	647	9	ADW45208	Adw45208 K. lactis

ALIGNMENTS

RESULT 1
ADF16526
ID ADF16526 standard; protein; 662 AA.
XX
AC ADF16526;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human albumin therapeutic fusion protein SeqID1623.
XX
KW albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human.
XX
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003060071-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040891.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.

PA (PRIN-) PRINCIPIA PHARM CORP.
XX
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598517/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Example 4; SEQ ID NO 1623; 24pp; English.
XX
CC This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is the amino acid sequence of a
CC novel full-length human albumin therapeutic fusion protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ Sequence 662 AA;

Query Match 100.0%; Score 3417; DB 7; Length 662;
Best Local Similarity 100.0%; Pred. No. 3.2e-283;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 77

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 137

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 197

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 257

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 317

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 377

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 437

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 497

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 557

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 617

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        618 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 662
```

RESULT 2
ADH21814
ID ADH21814 standard; protein; 662 AA.
XX
AC ADH21814;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:611.
XX
KW Fusion protein; human serum albumin; HSA; therapeutic protein;
KW shelf-life; in vitro biological activity; in vivo biological activity;
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW anorectic; ophthalmological; gene therapy.
XX
OS Synthetic.
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003059934-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040892.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598501/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Disclosure; SEQ ID NO 611; 1086pp; English.
XX
CC The invention relates to fusion proteins comprising human serum albumin
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC antibody or peptide or their variants or fragments. The therapeutic
CC protein may be fused to the N-terminus, the C-terminus or both termini of
CC albumin via a linker. The albumin component of the fusion proteins
CC prolongs the shelf-life and the in vitro and vivo biological activity of
CC the proteins compared with those of the corresponding therapeutic
CC proteins on their own. The invention also relates to nucleic acids
CC encoding albumin fusion proteins, vectors and host cells comprising an
CC albumin fusion protein nucleic acid, compositions and kits comprising an
CC albumin fusion protein, the method of extending the shelf-life of a
CC therapeutic protein by fusion with albumin, and the treatment of disease
CC using an albumin fusion protein. The albumin fusion proteins may be used
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-

```
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 662 AA;

  Query Match             100.0%;  Score 3417;  DB 7;  Length 662;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 77

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 137

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 197

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 257

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 317

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 377

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 437

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 497

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 557

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 617

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        618 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 662


RESULT 3
ADF16512
ID   ADF16512 standard; protein; 663 AA.
XX
AC   ADF16512;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1609.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1609; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 663 AA;

  Query Match             100.0%;  Score 3417;  DB 7;  Length 663;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 78

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 138
```

```
Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 198

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 258

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 318

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 378

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 438

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 498

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 558

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 618

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        619 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 663
```

```
RESULT 4
ADH21803
ID   ADH21803 standard; protein; 663 AA.
XX
AC   ADH21803;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:600.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Disclosure; SEQ ID NO 600; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 663 AA;
```

```
  Query Match             100.0%;  Score 3417;  DB 7;  Length 663;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 78

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 138

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 198

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 258

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 318

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 378

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 438
```

```
Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 498

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 558

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 618

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        619 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 663
```

RESULT 5
ADF16510
ID ADF16510 standard; protein; 664 AA.
XX
AC ADF16510;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human albumin therapeutic fusion protein SeqID1607.
XX
KW albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human.
XX
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003060071-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040891.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
PA (PRIN-) PRINCIPIA PHARM CORP.
XX
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598517/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Example 4; SEQ ID NO 1607; 24pp; English.
XX
CC This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is the amino acid sequence of a
CC novel full-length human albumin therapeutic fusion protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ Sequence 664 AA;

```
  Query Match             100.0%;  Score 3417;  DB 7;  Length 664;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 79

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 139

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 199

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 259

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 319

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 379

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 439

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 499

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 559

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 619

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        620 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 664
```

RESULT 6
ADH21801
ID ADH21801 standard; protein; 664 AA.
XX

```
AC   ADH21801;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G))x2 fusion protein, SEQ ID NO:598.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Disclosure; SEQ ID NO 598; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 664 AA;

  Query Match             100.0%;  Score 3417;  DB 7;  Length 664;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 79

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 139

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 199

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 259

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 319

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 379

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 439

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 499

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 559

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 619

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        620 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 664


RESULT 7
ADF16524
ID   ADF16524 standard; protein; 668 AA.
XX
AC   ADF16524;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1621.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
```

```
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1621; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 668 AA;

  Query Match             100.0%;  Score 3417;  DB 7;  Length 668;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 83

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 143

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 203

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 263

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 323

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 383

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 443

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 503

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 563

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 623

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        624 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 668


RESULT 8
ADH21812
ID   ADH21812 standard; protein; 668 AA.
XX
AC   ADH21812;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G))x2 fusion protein, SEQ ID NO:609.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
```

```
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Disclosure; SEQ ID NO 609; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 668 AA;

  Query Match             100.0%;  Score 3417;  DB 7;  Length 668;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 83

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 143

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 203

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 263

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 323

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 383

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 443

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 503

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 563

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 623

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        624 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 668


RESULT 9
ADF16144
ID   ADF16144 standard; protein; 669 AA.
XX
AC   ADF16144;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1231.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1231; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
```

```
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 669 AA;

  Query Match             100.0%;  Score 3417;  DB 7;  Length 669;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 84

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 144

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 204

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 264

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 324

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 384

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 444

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 504

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 564

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 624

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        625 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 669


RESULT 10
ADH21622
ID   ADH21622 standard; protein; 669 AA.
XX
AC   ADH21622;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:419.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Disclosure; SEQ ID NO 419; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 669 AA;

  Query Match             100.0%;  Score 3417;  DB 7;  Length 669;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 84

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 144

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 204

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 264

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 324

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 384

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 444

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 504

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 564

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 624

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
                 |||||||||||||||||||||||||||||||||||||||||||
Db        625 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 669


RESULT 11
ADF16193
ID   ADF16193 standard; protein; 674 AA.
XX
AC   ADF16193;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1280.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1280; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 674 AA;

  Query Match           100.0%;  Score 3417;  DB 7;  Length 674;
  Best Local Similarity  100.0%;  Pred. No. 3.3e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 149

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 209

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 269

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 329

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 389
```

```
Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 449

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 509

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 569

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 629

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        630 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674


RESULT 12
ADH21650
ID   ADH21650 standard; protein; 674 AA.
XX
AC   ADH21650;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G))x2 fusion protein, SEQ ID NO:447.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Disclosure; SEQ ID NO 447; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 674 AA;

  Query Match             100.0%;  Score 3417;  DB 7;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 149

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 209

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 269

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 329

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 389

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 449

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 509

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 569

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 629

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
```

```
                                     ||||||||||||||||||||||||||||||||||||||||||||
Db        630 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674


RESULT 13
ADW45202
ID   ADW45202 standard; protein; 674 AA.
XX
AC   ADW45202;
XX
DT   07-APR-2005  (first entry)
XX
DE   K. lactis killer toxin-GLP1-human serum albumin fusion protein - SEQ 206.
XX
KW   fusion protein; anti-HIV; gastrointestinal-gen.; antidiabetic; anorectic;
KW   nephrotropic; cardiant; cytostatic; neuroprotective; immunosuppressive;
KW   immune disorder; hematological disease; hyperproliferative disorder;
KW   renal disease; cardiovascular disease; cardiovascular-gen.;
KW   respiratory disorder; angiogenesis disorder; neurological disease;
KW   wound healing; vulnerary; endocrine disease; reproductive disorder;
KW   gynecological; infectious disease; antimicrobial;
KW   gastrointestinal disease; gene therapy; toxin; HSA; albumin;
KW   glucagon-like peptide 1; GLP1.
XX
OS   Homo sapiens.
OS   Kluyveromyces lactis.
OS   Chimeric.
XX
PN   WO2005003296-A2.
XX
PD   13-JAN-2005.
XX
PF   20-JAN-2004; 2004WO-US001369.
XX
PR   22-JAN-2003; 2003US-0441305P.
PR   11-MAR-2003; 2003US-0453201P.
PR   02-MAY-2003; 2003US-0467222P.
PR   23-MAY-2003; 2003US-0472816P.
PR   06-JUN-2003; 2003US-0476267P.
PR   24-SEP-2003; 2003US-0505172P.
PR   30-SEP-2003; 2003US-0506746P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Haseltine WA,  Rosen CA;
XX
DR   WPI; 2005-091786/10.
XX
PT   New albumin fusion protein for diagnosing, treating or preventing
PT   diseases such as HIV/AIDS, diabetes, obesity, heart disease or immune
PT   disorders comprises a therapeutic protein (e.g. CD4M33, GLP-2 or PACAP-
PT   27) and an albumin.
XX
PS   Example 13; SEQ ID NO 206; 884pp; English.
XX
CC   The invention relates to a novel albumin fusion protein comprising a
CC   therapeutic protein as listed in the specification in Table 1 and an
CC   albumin comprising a sequence of SEQ ID NO: 1, or a fragment or variant
CC   of SEQ ID NO: 1, where the fragment or variant has albumin activity and
CC   where the albumin activity is the ability to prolong the shelf life of
CC   the therapeutic protein compared to the shelf-life of the therapeutic
CC   protein in an unfused state. Human serum albumin (HSA, HA) is responsible
CC   for a significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion
CC   protein of the invention demonstrates anti-HIV, gastrointestinal-gen.,
CC   antidiabetic, anorectic, cardiant and immunosuppressive activities. The
CC   fusion protein may be useful for diagnosing, treating, preventing or
CC   ameliorating diseases, such as immune disorders, blood disorders,
CC   hyperproliferative disorders, renal disorders, cardiovascular disorders,
CC   respiratory disorders, angiogenesis-related disorders, neurological
CC   disorders, wound healing disorders, endocrine disorders, reproductive
CC   disorders, infectious disorders and gastrointestinal disorders, possibly
CC   with the use of gene therapy techniques. The current sequence is that of
CC   the Kluyveromyces lactis killer toxin-GLP1-human serum albumin fusion
CC   protein - SEQ 206 of the invention.
XX
SQ   Sequence 674 AA;

  Query Match             100.0%;  Score 3417;  DB 9;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 149

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 209

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 269

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 329

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 389

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 449

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 509

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 569

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 629

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        630 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674


RESULT 14
ADF16525
ID   ADF16525 standard; protein; 730 AA.
XX
AC   ADF16525;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1622.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
```

```
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1622; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 730 AA;

  Query Match             100.0%;  Score 3417;  DB 7;  Length 730;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-283;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 145

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 205

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 265

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 325

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 385

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 445

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 505

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 565

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 625

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 685

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        686 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 730


RESULT 15
ADH21813
ID   ADH21813 standard; protein; 730 AA.
XX
AC   ADH21813;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:610.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
```

PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598501/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Disclosure; SEQ ID NO 610; 1086pp; English.
XX
CC The invention relates to fusion proteins comprising human serum albumin
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC antibody or peptide or their variants or fragments. The therapeutic
CC protein may be fused to the N-terminus, the C-terminus or both termini of
CC albumin via a linker. The albumin component of the fusion proteins
CC prolongs the shelf-life and the in vitro and vivo biological activity of
CC the proteins compared with those of the corresponding therapeutic
CC proteins on their own. The invention also relates to nucleic acids
CC encoding albumin fusion proteins, vectors and host cells comprising an
CC albumin fusion protein nucleic acid, compositions and kits comprising an
CC albumin fusion protein, the method of extending the shelf-life of a
CC therapeutic protein by fusion with albumin, and the treatment of disease
CC using an albumin fusion protein. The albumin fusion proteins may be used
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC related conditions. Specifically the albumin fusion proteins may be used
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC (especially neuropathy), retinopathy, cardiovascular disorders
CC (especially heart disease, renal disorders and obesity. The proteins may
CC also be used in a method of maintaining a basal glucose level in a
CC patient and in a method for losing weight. The present sequence is
CC related to the invention.
XX
SQ Sequence 730 AA;

Query Match 100.0%; Score 3417; DB 7; Length 730;
Best Local Similarity 100.0%; Pred. No. 3.7e-283;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 145

Qy   61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  146 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 205

Qy  121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  206 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 265

Qy  181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  266 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 325

Qy  241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  326 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 385

Qy  301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  386 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 445

Qy  361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  446 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 505

Qy  421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  506 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 565

Qy  481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  566 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 625

Qy  541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  626 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 685

Qy  601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
          |||||||||||||||||||||||||||||||||||||||||||||
Db  686 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 730
```

Search completed: April 19, 2006, 12:02:33
Job time : 155.004 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:02:52 ; Search time 29.3182 Seconds
(without alignments)
2116.769 Million cell updates/sec

Title: US-10-775-180-447_COPY_30_674
Perfect score: 3417
Sequence: 1 HGEGTFTSDVSSYLEGQAAK.....TCFAEEGKKLVAASQAALGL 645

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3108	91.0	609	1	ABHUS	serum albumin prec
2	2947	86.2	600	2	A47391	serum albumin prec
3	2627	76.9	608	2	S57632	serum albumin prec
4	2481.5	72.6	607	1	ABHOS	serum albumin prec
5	2451.5	71.7	607	1	ABBOS	serum albumin prec
6	2437.5	71.3	607	1	ABSHS	serum albumin prec
7	2431	71.1	608	1	ABRTS	serum albumin prec
8	2416.5	70.7	605	1	ABPGS	serum albumin prec
9	2387.5	69.9	609	2	JC5838	albumin - Mongolia
10	1861	54.5	453	2	A05139	serum albumin - mo
11	1562	45.7	615	1	ABCHS	serum albumin prec
12	1260.5	36.9	609	2	JC4258	alpha-fetoprotein
13	1256.5	36.8	609	1	FPHU	alpha-fetoprotein
14	1249.5	36.6	609	1	FPGO	alpha-fetoprotein
15	1207.5	35.3	607	1	ABXL72	74K albumin precur
16	1181.5	34.6	265	2	I46986	albumin - dog (fra
17	1175.5	34.4	608	1	ABXL68	68K serum albumin
18	1084	31.7	605	1	FPMS	alpha-fetoprotein
19	1067	31.2	611	1	FPRT	alpha-fetoprotein
20	1055	30.9	599	1	A54906	afamin precursor -
21	932.5	27.3	608	2	A53195	afamin precursor -
22	930	27.2	614	2	S59517	serum albumin prec
23	751.5	22.0	608	1	ABONS1	serum albumin 1 pr
24	746.5	21.8	608	1	ABONS2	serum albumin 2 pr
25	699	20.5	382	2	A37253	serum albumin - bu
26	440.5	12.9	1423	1	S27941	serum albumin - se
27	401	11.7	474	1	VYHUD	vitamin D-binding
28	400	11.7	476	1	VYRTD	vitamin D-binding
29	387	11.3	472	1	A35327	vitamin D-binding
30	227.5	6.7	180	1	GCRT	glucagon precursor
31	227.5	6.7	180	2	A57294	glucagon precursor
32	225.5	6.6	180	1	GCBO	glucagon precursor
33	225.5	6.6	180	1	GCHY	glucagon precursor
34	225.5	6.6	180	1	GCHU	glucagon precursor
35	224.5	6.6	180	1	GCGP	glucagon precursor
36	222.5	6.5	158	1	GCPG	glucagon precursor
37	214.5	6.3	180	1	GCRTDU	glucagon precursor
38	213	6.2	206	2	I51301	proglucagon - chic
39	209	6.1	101	1	GCFGB	glucagon precursor
40	204.5	6.0	151	1	GCCH	glucagon precursor
41	190.5	5.6	122	1	GCAF2	glucagon 2 precurs
42	188.5	5.5	63	1	GCIDC	glucagon precursor
43	188.5	5.5	178	2	I51058	glucagon I precurs
44	188	5.5	72	1	GCGXA	glucagon precursor
45	184	5.4	1819	2	A71928	cag island protein

ALIGNMENTS

RESULT 1
ABHUS
serum albumin precursor [validated] - human
N;Alternate names: preproalbumin
N;Contains: kinetensin
C;Species: Homo sapiens (man)
C;Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C;Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S3
R;Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu
Nucleic Acids Res. 9, 6103-6114, 1981
A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia col
A;Reference number: A93743; MUID:82081882; PMID:6171778
A;Accession: A93743
A;Molecule type: mRNA
A;Residues: 1-419,'K',421-609 <LAW>
A;Cross-references: UNIPROT:P02768; UNIPARC:UPI000002CE3A; EMBL:V00495; GB:J00078; GB:L
R;Dugaiczyk, A.; Law, S.W.; Dennison, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A;Reference number: A93936; MUID:82105994; PMID:6275391
A;Accession: A93936
A;Molecule type: mRNA
A;Residues: 1-120,'G',122-609 <DUG>
A;Cross-references: UNIPARC:UPI0000156BE8; EMBL:V00494; NID:g28589; PIDN:CAA23753.1; PI
R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A;Reference number: I39427; MUID:86140099; PMID:2419329
A;Accession: I39427
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-26 <URA>
A;Cross-references: UNIPARC:UPI000002BD5F; GB:M13075; NID:g178330; PIDN:AAA51688.1; PID
R;Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family
A;Reference number: I59286; MUID:94181575; PMID:8134387
A;Accession: I59286
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 282-290,'KSRFDLQ' <WAT>
A;Cross-references: UNIPARC:UPI000011F7AF; GB:S69192; NID:g546032; PIDN:AAB30282.1; PID
A;Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R;Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-
A;Reference number: I59313; MUID:94294404; PMID:8022807
A;Accession: I59313
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 589-590,'ALPRRVKNLLLQVKLP' <MAD>
A;Cross-references: UNIPARC:UPI0000072EC4; GB:S70799; NID:g547231; PIDN:AAB31177.1; PID

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants
R;Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A;Reference number: G08292
A;Accession: G01747
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-120,'G',122-455 <MEN>
A;Cross-references: UNIPARC:UPI000016A1A8; EMBL:U22961; NID:g763428; PIDN:AAA64922.1; PI
R;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
A;Reference number: S55314; MUID:95275251; PMID:7755581
A;Accession: S55314
A;Molecule type: protein
A;Residues: 19-27 <LED>
A;Cross-references: UNIPARC:UPI00001743FA
R;Meloun, B.; Moravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A;Title: Complete amino acid sequence of human serum albumin.
A;Reference number: A91420; MUID:76187907; PMID:1225573
A;Accession: A91420
A;Molecule type: protein
A;Residues: 25-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',395-
A;Cross-references: UNIPARC:UPI00001743FB
R;Roehr, U.; Spiteller, G.; Tripier, D.
Justus Liebigs Ann. Chem. 9, 881-884, 1988
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from ur
A;Reference number: S06422
A;Note: this paper is in German, with an English abstract
A;Accession: S06422
A;Molecule type: protein
A;Residues: 25-48 <ROE>
A;Cross-references: UNIPARC:UPI0000052CDA
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A;Title: Mass spectrometric identification of modifications to human serum albumin treat
A;Reference number: S36882; MUID:93384321; PMID:8373198
A;Accession: S36882
A;Molecule type: protein
A;Residues: 45-67;141-160;311-337;469-490;570-581 <FIN>
A;Cross-references: UNIPARC:UPI00000423AC; UNIPARC:UPI00001743FC; UNIPARC:UPI00001743FD;
R;Kausler, E.; Spiteller, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A;Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol
A;Reference number: S17599; MUID:92126241; PMID:1772598
A;Accession: S17599
A;Molecule type: protein
A;Residues: 25-54;354-357;431-447 <KAU>
A;Cross-references: UNIPARC:UPI0000174400; UNIPARC:UPI0000174401; UNIPARC:UPI0000174402
A;Note: 49-Leu was also found
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: A45800
A;Molecule type: protein
A;Residues: 166-173 <CAR>
A;Cross-references: UNIPARC:UPI000004A560
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre
A;Reference number: A03239; MUID:86242180; PMID:3087352
A;Accession: A03239
A;Molecule type: protein
A;Residues: 166-173,'L' <MOG>
A;Cross-references: UNIPARC:UPI00000351D2
R;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A;Title: Mutations in genetic variants of human serum albumin found in Italy.
A;Reference number: A38255; MUID:91062352; PMID:2247440
A;Accession: C38255
A;Molecule type: protein
A;Residues: 76-111 <GAL1>
A;Cross-references: UNIPARC:UPI0000174403
A;Accession: B38255
A;Molecule type: protein
A;Residues: 82-105,'K',107-110 <GAL2>
A;Cross-references: UNIPARC:UPI0000174403
A;Note: this variant is designated albumin Vibo Valentia
A;Accession: A38255
A;Molecule type: protein
A;Residues: 76-83,'K',85-106 <GAL3>
A;Cross-references: UNIPARC:UPI0000174405
A;Note: this variant is designated albumin Torino
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin Her
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: protein
A;Residues: 255-263,'E',265-281 <MIN1>
A;Cross-references: UNIPARC:UPI0000174406
A;Note: this variant is designated albumin Herborn
R;Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,
Biochim. Biophys. Acta 1119, 232-238, 1992
A;Title: Two alloalbumins with identical electrophoretic mobility are produced by differ
A;Reference number: S21078; MUID:92190239; PMID:1347703
A;Accession: S21078
A;Molecule type: protein
A;Residues: 354-356,'K',358-378 <MIN2>
A;Cross-references: UNIPARC:UPI0000174407
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,
R;He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A;Title: Atomic structure and chemistry of human serum albumin.
A;Reference number: A46756; MUID:92334427; PMID:1630489
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,
A;Reference number: A94442
A;Contents: annotation; three-dimensional structure and disulfide bonds
R;Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A;Title: Disulfide bonds in human serum albumin.
A;Reference number: A90930
A;Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A;Reference number: A90299; MUID:78186630; PMID:656055
A;Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjoholm, I., eds., 11-20,
A;Title: Serum albumin: conformation and active sites.
A;Reference number: A94408
A;Contents: annotation; binding sites
R;Harper, M.E.; Dugaiczyk, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes
A;Reference number: A90028; MUID:83279982; PMID:6192711
A;Contents: annotation; gene position
R;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicyclic acid.
A;Reference number: A46755; MUID:76257808; PMID:955075
A;Contents: annotation
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicyclic acid
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phospha
A;Reference number: A56294; MUID:92183881; PMID:1544460
A;Contents: annotation
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p
atase activity
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized

lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
C;Comment: A large number of variants of human serum albumin have been described.
C;Genetics:
A;Gene: GDB:ALB
A;Cross-references: GDB:118990; OMIM:103600
A;Map position: 4q11-4q13

Query Match 91.0%; Score 3108; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.8e-196;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy         120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy         180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy         240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy         300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy         360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy         420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy         480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy         540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy         600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
               |||||||||||||||||||||||||||||||||||||||||||||
Db         564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

RESULT 2
A47391
serum albumin precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A47391
R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A;Title: cDNA and protein sequence of polymorphic macaque albumins that differ in biliru
A;Reference number: A47391; MUID:93211971; PMID:8460152
A;Contents: B/B homozygote
A;Accession: A47391
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-600 <WAT>
A;Cross-references: UNIPROT:Q28522; UNIPARC:UPI000001257C4; GB:M90463; NID:g342294; PIDN:
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)
C;Superfamily: serum albumin; serum albumin repeat homology
F;21-194/Domain: serum albumin repeat homology <SA1>
F;213-386/Domain: serum albumin repeat homology <SA2>
F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 86.2%; Score 2947; DB 2; Length 600;
Best Local Similarity 93.5%; Pred. No. 6.5e-186;
Matches 546; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

```
Qy          60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
               || |||||||||||||||:|| ||| :||:|||||||||:||||||||||||||||||||||
Db          16 RDTHKSEVAHRFKDLGEEHFKGLVLVAFSQYLQQCPFEEHVKLVNEVTEFAKTCVADESA 75

Qy         120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          76 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPPLVRPE 135

Qy         180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
               |||||||||||| |||||||||:|||||||||||||| |||||| ||||||||||||||
Db         136 VDVMCTAFHDNEATFLKKYLYEVARRHPYFYAPELLFFAARYKAAFAECCQAADKAACLL 195

Qy         240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
               ||||||||||||||||||||||||||:||||||||||||:||||||||||||||||||||
Db         196 PKLDELRDEGKASSAKQRLKCASLQKFGDRAFKAWAVARLSQKFPKAEFAEVSKLVTDLT 255

Qy         300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
               ||||||||||||||||||||||||:||||||||||||:|||||||||:|||||||||||||
Db         256 KVHTECCHGDLLECADDRADLAKYMCENQDSISSKLKECCDKPLLEKSHCLAEVENDEMP 315

Qy         360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
               |||||||||:|||||||||||||||||||||||||||||||||:||||||| || ||||||
Db         316 ADLPSLAADYVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVMLLLRLAKAYEATLEK 375

Qy         420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
               ||||||||||||||||||||:|||||||||:|||||||||||||||||||||||||||||||
Db         376 CCAAADPHECYAKVFDEFQPLVEEPQNLVKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 435

Qy         480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
               ||||||||||||||:|||| ||||||||||||||||:|||||||||::|||||||||||||
Db         436 TPTLVEVSRNLGKVGAKCCKLPEAKRMPCAEDYLSVVLNRLCVLHEKTPVSEKVTKCCTE 495

Qy         540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
               |||||||||||||:|| |||| |||||||||||:|||||||:|:||||||||||||||||
Db         496 SLVNRRPCFSALELDEAYVPKAFNAETFTFHADMCTLSEKEKQVKKQTALVELVKHKPKA 555

Qy         600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 643
               |||||| |||:|||||||||||||| |||||| | ||||||||
Db         556 TKEQLKGVMDNFAAFVEKCCKADDKEACFAEEGPKFVAASQAAL 599
```

RESULT 3
S57632
serum albumin precursor - cat
C;Species: Felis silvestris catus (domestic cat)
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: JC4660; S57632
R;Hilger, C.; Grigioni, F.; Hentges, F.
Gene 169, 295-296, 1996
A;Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A;Reference number: JC4660; MUID:96194824; PMID:8647469
A;Accession: JC4660
A;Molecule type: mRNA
A;Residues: 1-608 <HI2>
A;Cross-references: UNIPROT:P49064; UNIPARC:UPI000001257C2; EMBL:X84842; NID:g886484; PI
A;Experimental source: liver
C;Comment: This protein is the major protein component in plasma. It functions as a mul
ein has 35 conserved cysteine residues.
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: liver; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRP>
F;25-608/Product: serum albumin #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>

```
  Query Match             76.9%;  Score 2627;  DB 2;  Length 608;
  Best Local Similarity   80.1%;  Pred. No. 6.7e-165;
  Matches  483;  Conservative   53;  Mismatches   57;  Indels   10;  Gaps    1;

Qy         41 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 100
              |:|  |           |  |:||:||:|||| |||||:|: |||:||:|||||||||||
Db         15 SAYSRG----------VTRREAHQSEIAHRFNDLGEEHFRGLVLVAFSQYLQQCPFEDHV 64

Qy        101 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 160
              |||||||||| ||||:|| ||:|||| | |||||||:||: |||||||  |:||||||
Db         65 KLVNEVTEFAKGCVADQSAANCEKSLHELLGDKLCTVASLRDKYGEMADCCEKKEPERNE 124

Qy        161 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 220
              ||||||||||   :|| || | ||||||:||: || ||||||||||||||||||::|:
Db        125 CFLQHKDDNPGFGQLVTPEADAMCTAFHENEQRFLGKYLYEIARRHPYFYAPELLYYAEE 184

Qy        221 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 280
              ||  |||||:|||||||| ||:| ||::  |||||:|||||||||||||||||:|||||
Db        185 YKGVFTECCEAADKAACLTPKVDALREKVLASSAKERLKCASLQKFGERAFKAWSVARLS 244

Qy        281 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 340
              |:|||||||:|||||| |:| |||||||||||||||||||||||||||:||||||
Db        245 QKFPKAEFAEISKLVTDLAKIHKECCHGDLLECADDRADLAKYICENQDSISTKLKECCG 304

Qy        341 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 400
              ||:|||||||:||| ||:||||| || |||| |:||||| ||||||| |||||:||||:
Db        305 KPVLEKSHCISEVERDELPADLPPLAVDFVEDKEVCKNYQEAKDVFLGTFLYEYSRRHPE 364

Qy        401 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 460
              ||| ||||||| || |||||||  ||  ||| |||||||||||| ||:| ||||||:|||
Db        365 YSVSLLLRLAKEYEATLEKCCATDDPPACYAHVFDEFKPLVEEPHNLVKTNCELFEKLGE 424

Qy        461 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 520
              | |||||||||||||||||||||||:|||||||| ||||:|: ||||||||||||:|
Db        425 YGFQNALLVRYTKKVPQVSTPTLVEVSRSLGKVGSKCCTHPEAERLSCAEDYLSVVLNRL 484

Qy        521 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 580
              ||||||||||:|||||||||||||||||:||||||||||:|||||||||:||| | |
Db        485 CVLHEKTPVSERVTKCCTESLVNRRPCFSALQVDETYVPKEFSAETFTFHADLCTLPEAE 544

Qy        581 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 640
              :|||||:|||||:|||||||:|||| || || :||:||| |:||| |||||| ||||:|
Db        545 KQIKKQSALVELLKHKPKATEEQLKTVMGDFGSFVDKCCAAEDKEACFAEEGPKLVAAAQ 604

Qy        641 AAL 643
              |||
Db        605 AAL 607


RESULT 4
ABHOS
serum albumin precursor - horse
C;Species: Equus caballus (domestic horse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S34053
R;Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A;Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A;Reference number: S34053; MUID:93345495; PMID:8344282
A;Accession: S34053
A;Molecule type: mRNA
A;Residues: 1-607 <HOA>
A;Cross-references: UNIPROT:P35747; UNIPARC:UPI00001257C3; GB:X74045; NID:g399671; PIDN:
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F;263/Binding site: bilirubin (Lys) #status predicted

  Query Match             72.6%;  Score 2481.5;  DB 1;  Length 607;
  Best Local Similarity   74.6%;  Pred. No. 2.4e-155;
  Matches  450;  Conservative   70;  Mismatches   72;  Indels   11;  Gaps    2;

Qy         41 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 100
              |:|  |           |  || ||||:|||| ||||::|| |||:||:||||||||||
Db         15 SAYSRG----------VLRRDTHKSEIAHRFNDLGEKHFKGLVLVAFSQYLQQCPFEDHV 64

Qy        101 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 160
              |||||||||| | |||||||||||||||||||||||| ||||:|||| ||||||||
Db         65 KLVNEVTEFAKKCAADESAENCDKSLHTLFGDKLCTVATLRATYGELADCCEKQEPERNE 124

Qy        161 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 220
              ||| ||||:||||:| :|| |  | || :: : || |||||:|||||||| ||||| |:
Db        125 CFLTHKDDHPNLPKL-KPEPDAQCAAFQEDPDKFLGKYLYEVARRHPYFYGPELLFHAEE 183

Qy        221 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 280
              ||| ||||| | || |||:|||| |::    ||||:||||:| | ||||| |||:|||||
Db        184 YKADFTECCPADDKLACLIPKLDALKERILLSSAKERLKCSSFQNFGERAVKAWSVARLS 243

Qy        281 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 340
              |:||||:||||||:|||||| |||||||||||||||||||||||:||||| ||| ||:
Db        244 QKFPKADFAEVSKIVTDLTKVHKECCHGDLLECADDRADLAKYICEHQDSISGKLKACCD 303

Qy        341 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 400
              ||||:|||||||: |::|:|||:|||| | |::||:| :|||||| ||||||||||||
Db        304 KPLLQKSHCIAEVKEDDLPSDLPALAADFAEDKEICKHYKDAKDVFLGTFLYEYSRRHPD 363

Qy        401 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 460
              ||| ||||:|||| |||||| |||   ||  |||:| |||||::|:|:||:|||::||
Db        364 YSVSLLLRIAKTYEATLEKCCAEADPPACYRTVFDQFTPLVEEPKSLVKKNCDLFEEVGE 423

Qy        461 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 520
              | |||||:|||||| ||||||||||: | ||||||:||| ||::|:||:|::|:: ||:|
Db        424 YDFQNALIVRYTKKAPQVSTPTLVEIGRTLGKVGSRCCKLPESERLPCSENHLALALNRL 483

Qy        521 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 580
              ||||||||||:::|||||:||  |||||||||:|| |||||| |||||||||||| | |
Db        484 CVLHEKTPVSEKITKCCTDSLAERRPCFSALELDEGYVPKEFKAETFTFHADICTLPEDE 543

Qy        581 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 640
              :|||||:|| |||||||||||||| |: :|:||| |||  :||| |||||| ||||:||
Db        544 KQIKKQSALAELVKHKPKATKEQLKTVLGNFSAFVAKCCGREDKEACFAEEGPKLVASSQ 603

Qy        641 AAL 643
              ||
Db        604 LAL 606


RESULT 5
ABBOS
serum albumin precursor [validated] - bovine
N;Alternate names: 67K protein; preproalbumin
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94
R;Holowachuk, E.W.; Stoltenborg, J.K.; Reed, R.G.; Peters Jr., T.
submitted to the EMBL Data Library, August 1991
A;Description: Bovine serum albumin: cDNA sequence and expression.
A;Reference number: A38885
A;Accession: A38885
A;Molecule type: mRNA
A;Residues: 1-607 <HOL>
A;Cross-references: UNIPROT:P04277; UNIPARC:UPI0000174408; EMBL:M73215
R;Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
A;Title: Rapid confirmation and revision of the primary structure of bovine serum albumi
A;Reference number: A36401; MUID:91083649; PMID:2260975
A;Accession: A36401
```

A;Molecule type: protein
A;Residues: 25-41,'H',43-189,'E',191-213,'T',215-323,'D',325-393,'TS',396-607 <HIR>
A;Cross-references: UNIPARC:UPI0000174409
R;MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-485, 1979
A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A;Reference number: A91258; MUID:80024278; PMID:488109
A;Accession: A91258
A;Molecule type: protein
A;Residues: 1-32 <MAG>
A;Cross-references: UNIPARC:UPI000017440A
R;Hsieh, J.C.; Lin, F.P.; Tam, M.F.
Anal. Biochem. 170, 1-8, 1988
A;Title: Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing g
A;Reference number: A60808; MUID:88267456; PMID:3389500
A;Accession: B60808
A;Molecule type: protein
A;Residues: 25-41 <HSI>
A;Cross-references: UNIPARC:UPI000017440B
R;Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albu
A;Reference number: S10780; MUID:90336641; PMID:2379503
A;Accession: S10780
A;Molecule type: protein
A;Residues: 25-41,'H',43-57,59-64 <STR>
A;Cross-references: UNIPARC:UPI000017440C
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: D45800
A;Molecule type: protein
A;Residues: 163-172 <CAR>
A;Cross-references: UNIPARC:UPI000017440D
R;Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.
J. Biol. Chem. 262, 5968-5973, 1987
A;Title: Structure of a biologically active neurotensin-related peptide obtained from pe
A;Reference number: A26693; MUID:87194805; PMID:2437111
A;Accession: A26693
A;Molecule type: protein
A;Residues: 165-172,'L' <CA2>
A;Cross-references: UNIPARC:UPI00000351D2
R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.
Biochem. J. 191, 867-868, 1980
A;Title: Sequence of residues 400-403 of bovine serum albumin.
A;Reference number: A90309; MUID:82023364; PMID:7283978
A;Accession: A90309
A;Molecule type: protein
A;Residues: 402-433 <REE>
A;Cross-references: UNIPARC:UPI000017440E
R;Brown, J.R.
Fed. Proc. 34, 591, 1975
A;Title: Structure of bovine serum albumin.
A;Reference number: A91458
A;Accession: A91458
A;Molecule type: protein
A;Residues: 25-41,'H',43-117,'EQ',120-179,181-189,'E',191-194,'A',196-213,'T',215-288,'E
A;Cross-references: UNIPARC:UPI000017440F; UNIPARC:UPI0000174410
R;Brown, J.R.
submitted to the Atlas, April 1975
A;Reference number: A94551
A;Accession: A94551
A;Molecule type: protein
A;Residues: 190-195 <BR2>
A;Cross-references: UNIPARC:UPI0000174411
R;Brown, J.R.
Fed. Proc. 33, 1389, 1974
A;Reference number: A91457
A;Contents: annotation; disulfide bonds
R;Werlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC

A;Reference number: S55232; MUID:95031935; PMID:7945219
A;Accession: S55232
A;Status: preliminary
A;Molecule type: protein
A;Residues: 529-536;569-572 <WER>
A;Cross-references: UNIPARC:UPI0000174412; UNIPARC:UPI0000174413
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; copper binding; duplication; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-607/Product: serum albumin #status experimental <MPT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4

Query Match 71.7%; Score 2451.5; DB 1; Length 607;
Best Local Similarity 75.7%; Pred. No. 2.2e-153;
Matches 442; Conservative 71; Mismatches 70; Indels 1; Gaps 1;

```
Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              || ||||:|||||||||| || ||||||:||||||||::|||||||:|||||||||||
Db         24 RDTHKSEIAHRFKDLGEEQFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFAKTCVADESH 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
               |:|||||||||:|| ||:||||||:||||| |||||||||||| ||||:|:||:| :|:
Db         84 AGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKL-KPD 142

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
               : :|  |  :|: |  ||||||||||||||||||::| :|   | :|||| || ||||
Db        143 PNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQDCCQAEDKGACLL 202

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||:: :|::  ||||:|||:|||:||||||| |||:||||||:|||||| ||:|||||||
Db        203 PKIETMREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKAEFVEVTKLVTDLT 262

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||| ||||||||||||||||||||:|||:|||||||||:|||||||||||||| | :|
Db        263 KVHKECCHGDLLECADDRADLAKYICDNQDTISSKLKECCDKPLLEKSHCIAEVEKDAIP 322

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
               :|| | ||| | ||||||| |||| ||| |||||:||||:|:| :|||||| || |||:
Db        323 ENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYEYSRRHPEYAVSVLLRLAKEYEATLEE 382

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              |||  ||| ||: ||| | ||:|||||||||: ||:|||| |||||:||||:|||||||
Db        383 CCAKDDPHACYSTVFDKLKHLVDEPQNLIKQNCDQFEKLGEYGFQNALIVRYTRKVPQVS 442

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              |||||||||:|||||::||  ||::|||| |||||::||:|||||||||||::|||||||
Db        443 TPTLVEVSRSLGKVGTRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTE 502

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              |||||||||||  ||||||| |: : |||||||||| : |:|||||||||||:||||||
Db        503 SLVNRRPCFSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQTALVELLKHKPKA 562

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 643
              |:|||| ||::| |||:||| ||||| ||| || ||| ::| ||
Db        563 TEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLVVSTQTAL 606
```

RESULT 6
ABSHS
serum albumin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S06936
R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A;Reference number: S06936; MUID:90098888; PMID:2602160
A;Accession: S06936

```
A;Molecule type: mRNA
A;Residues: 1-607 <BRO>
A;Cross-references: UNIPROT:P14639; UNIPARC:UPI00001257CB; EMBL:X17055; NID:g1386; PIDN:
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F;263/Binding site: bilirubin (Lys) #status predicted

  Query Match             71.3%;  Score 2437.5;  DB 1;  Length 607;
  Best Local Similarity   75.0%;  Pred. No. 1.8e-152;
  Matches  438;  Conservative   73;  Mismatches   72;  Indels    1;  Gaps    1;

Qy          60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
               || ||||:|||| ||||||: ||||||:||||||::||||| |:||||||||||||
Db          24 RDTHKSEIAHRFNDLGEENFQGLVLIAFSQYLQQCPFDEHVKLVKELTEFAKTCVADESH 83

Qy         120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
                 |||||||||||:|| |||||||||:||||| |||||||||| |||||:|:||:| :||
Db          84 AGCDKSLHTLFGDELCKVATLRETYGDMADCCEKQEPERNECFLNHKDDSPDLPKL-KPE 142

Qy         180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
               | :|  |  :|: |  |||||:||||||||||||::| :|   | ||||| || ||||
Db         143 PDTLCAEFKADEKKFWGKYLYEVARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLL 202

Qy         240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
               ||:| :|::  ||||:|||:|||:||||||| |||:||||||:||||:| :|:|:|||||
Db         203 PKIDAMREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKADFTDVTKIVTDLT 262

Qy         300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
               ||| ||||||||||||||||||::|::|||||||:||:|||||||||: | :|
Db         263 KVHKECCHGDLLECADDRADLAKYICDHQDALSSKLKECCDKPVLEKSHCIAEVDKDAVP 322

Qy         360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
                :|| | ||| | |:|||| |||||||| |||||:||||:|:| :|||||| || |||
Db         323 ENLPPLTADFAEDKEVCKNYQEAKDVFLGSFLYEYSRRHPEYAVSVLLRLAKEYEATLED 382

Qy         420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
               |||  ||| ||| |||: | ||:|||||||:||||||: ||| |||||:||||:| ||||
Db         383 CCAKEDPHACYATVFDKLKHLVDEPQNLIKKNCELFEKHGEYGFQNALIVRYTRKAPQVS 442

Qy         480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
               ||||||:||:|||||:|||  ||::|||| |||||::||:|||||||||||::|||||||
Db         443 TPTLVEISRSLGKVGTKCCAKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTE 502

Qy         540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
               |||||||||| | :||||||| |: : |||||||||| : |:|||||||||:||||||
Db         503 SLVNRRPCFSDLTLDETYVPKPFDEKFFTFHADICTLPDTEKQIKKQTALVELLKHKPKA 562

Qy         600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 643
               | |||| ||::| |||:||| ||||| ||  || ||||::||||
Db         563 TDEQLKTVMENFVAFVDKCCAADDKEGCFVLEGPKLVASTQAAL 606


RESULT 7
ABRTS
serum albumin precursor - rat
N;Alternate names: preproalbumin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 09-Jul-2004
C;Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R;Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A;Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A;Reference number: A93872; MUID:81223722; PMID:7017712
A;Accession: A93872
A;Molecule type: mRNA
A;Residues: 1-608 <SAR>
A;Cross-references: UNIPROT:P02770; UNIPARC:UPI00001257CA; GB:V01222; GB:J00698; NID:g55
R;Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A;Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analys
A;Reference number: A92211; MUID:77249657; PMID:893447
A;Note: cleavages during protein maturation
A;Accession: A92211
A;Molecule type: protein
A;Residues: 1-38 <STR>
A;Cross-references: UNIPARC:UPI0000174416
R;Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A;Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavag
A;Reference number: A91946; MUID:78109429; PMID:564345
A;Accession: A91946
A;Molecule type: protein
A;Residues: 25-222 <IS1>
A;Cross-references: UNIPARC:UPI0000174417
R;Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A;Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino a
A;Reference number: A91940; MUID:76260153; PMID:956149
A;Accession: A91940
A;Molecule type: protein
A;Residues: 223-288;572-608 <IS2>
A;Cross-references: UNIPARC:UPI0000174418; UNIPARC:UPI0000174419
A;Note: 262-Leu was also found
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.
A;Reference number: A90758; MUID:79001617; PMID:80265
A;Contents: annotation; copper binding
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: C45800
A;Status: preliminary
A;Molecule type: protein
A;Residues: 166-173 <CAR>
A;Cross-references: UNIPARC:UPI000017441A
R;Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A;Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved
A;Reference number: I57621; MUID:87286876; PMID:3475566
A;Accession: I57621
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: UNIPARC:UPI000011E8B8; GB:M16825; NID:g202828; PIDN:AAA40712.1; PID:
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-608/Product: serum albumin #status experimental <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status experimental
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4

  Query Match             71.1%;  Score 2431;  DB 1;  Length 608;
  Best Local Similarity   73.5%;  Pred. No. 4.9e-152;
  Matches  429;  Conservative   82;  Mismatches   73;  Indels    0;  Gaps    0;

Qy          60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
               |:|||||:|||||||||::|| |||||||:||||:||:|:|:||| |||:||||||||:|
Db          24 REAHKSEIAHRFKDLGEQHFKGLVLIAFSQYLQKCPYEEHIKLVQEVTDFAKTCVADENA 83

Qy         120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
```

```
Db         84 ENCDKSIHTLFGDKLCAIPKLRDNYGELADCCAKQEPERNECFLQHKDDNPNLPPFQRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239

Db        144 AEAMCTSFQENPTSFLGHYLHEVARRHPYFYAPELLYYAEKYNEVLTQCCTESDKAACLT 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299

Db        204 PKLDAVKEKALVAAVRQRMKCSSMQRFGERAFKAWAVARMSQRFPNAEFAEITKLATDVT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359

Db        264 KINKECCHGDLLECADDRAELAKYMCENQATISSKLQACCDKPVLQKSQCLAETEHDNIP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419

Db        324 ADLPSIAADFVEDKEVCKNYAEAKDVFLGTFLYEYSRRHPDYSVSLLLRLAKKYEATLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479

Db        384 CCAEGDPPACYGTVLAEFQPLVEEPKNLVKTNCELYEKLGEYGFQNAVLVRYTQKAPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539

Db        444 TPTLVEAARNLGRVGTKCCTLPEAQRLPCVEDYLSAILNRLCVLHEKTPVSEKVTKCCSG 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599

Db        504 SLVERRPCFSALTVDETYVPKEFKAETFTFHSDICTLPDKEKQIKKQTALAELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 643

Db        564 TEDQLKTVMGDFAQFVDKCCKAADKDNCFATEGPNLVARSKEAL 607


RESULT 8
ABPGS
serum albumin precursor - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S01382; A61006
R;Weinstock, J.; Baldwin, G.S.
Nucleic Acids Res. 16, 9045, 1988
A;Title: Nucleotide sequence of porcine liver albumin.
A;Reference number: S01382; MUID:89016582; PMID:3174440
A;Accession: S01382
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-605 <WEI>
A;Cross-references: UNIPROT:P08835; UNIPARC:UPI00001257C7; EMBL:X12422; NID:g1875; PIDN:
R;Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
J. Bone Miner. Res. 4, 235-241, 1989
A;Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral
A;Reference number: A61006; MUID:89269769; PMID:2728927
A;Accession: A61006
A;Molecule type: protein
A;Residues: 23-51,'X',53-54;'XXXGY',146,'E',148,'E',150-151,'XVN',155 <LIM>
A;Cross-references: UNIPARC:UPI0000174414; UNIPARC:UPI0000174415
A;Experimental source: dental enamel
A;Note: albumin and other serum proteins are also found in bone
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-16/Domain: signal sequence (fragment) #status predicted <SIG>
F;17-22/Domain: propeptide #status predicted <PRO>
F;23-605/Product: serum albumin #status predicted <MAT>
F;27-199/Domain: serum albumin repeat homology <SA1>
F;218-391/Domain: serum albumin repeat homology <SA2>
F;410-589/Domain: serum albumin repeat homology <SA3>
F;75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4
F;261/Binding site: bilirubin (Lys) #status predicted

  Query Match           70.7%;  Score 2416.5;  DB 1;  Length 605;
  Best Local Similarity 76.1%;  Pred. No. 4.4e-151;
  Matches  439;  Conservative   67;  Mismatches   70;  Indels    1;  Gaps    1;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119

Db         22 RDTYKSEIAHRFKDLGEQYFKGLVLIAFSQHLQQCPYEEHVKLVREVTEFAKTCVADESA 81

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179

Db         82 ENCDKSIHTLFGDKLCAIPSLREHYGDLADCCEKEEPERNECFLQHKNDNPDIPKL-KPD 140

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239

Db        141 PVALCADFQEDEQKFWGKYLYEIARRHPYFYAPELLYYAIIYKDVFSECCQAADKAACLL 200

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299

Db        201 PKIEHLREKVLTSAAKQRLKCASIQKFGERAFKAWSLARLSQRFPKADFTEISKIVTDLA 260

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359

Db        261 KVHKECCHGDLLECADDRADLAKYICENQDTISTKLKECCDKPLLEKSHCIAEAKRDELP 320

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419

Db        321 ADLNPLEHDFVEDKEVCKNYKEAKDVFLGTFLYEYSRRHPDYSVSLLLRIAKIYEATLED 380

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479

Db        381 CCAKEDPPACYATVFDKFQPLVDEPKNLIKQNCELFEKLGEYGFQNALIVRYTKKVPQVS 440

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539

Db        441 TPTLVEVARKLGLVGSRCCKRPEEERLSCAEDYLSLVLNRLCVLHEKTPVSEKVTKCCTE 500

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599

Db        501 SLVNRRPCFSALTPDETYKPKEFVEGTFTFHADLCTLPEDEKQIKKQTALVELLKHKPHA 560

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLV 636

Db        561 TEEQLRTVLGNFAAFVQKCCAAPDHEACFAVEGPKFV 597


RESULT 9
JC5838
albumin - Mongolian jird
C;Species: Meriones unguiculatus (Mongolian jird)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5838
R;Yoshida, K.; Seto-Ohshima, A.; Sinohara, H.
DNA Res. 4, 351-354, 1997
A;Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in th
A;Reference number: JC5838; MUID:98116663; PMID:9455485
A;Accession: JC5838
A;Molecule type: mRNA
A;Residues: 1-609 <YOS>
A;Cross-references: UNIPROT:O35090; UNIPARC:UPI00001257C5; DDBJ:AB006197; NID:g2317277;
A;Experimental source: liver
C;Superfamily: serum albumin; serum albumin repeat homology
F;222-395/Domain: serum albumin repeat homology <SA2>

  Query Match           69.9%;  Score 2387.5;  DB 2;  Length 609;
  Best Local Similarity 73.8%;  Pred. No. 3.5e-149;
  Matches  432;  Conservative   65;  Mismatches   87;  Indels    1;  Gaps    1;

Qy         60 RD-AHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADES 118

Db         24 RDAAHKSEIAHRYKDLGEKYFKGLVLYTFSQYLQKCSYEEHVKLVREVTDFASNCAKDES 83

Qy        119 AENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRP 178
```

```
Db         84 AENCDKSLHTLFGDKLCSLPNFGEKYAEMADCCAKQEPERNECFLQHKDDNPQLPPFKRA 143

Qy        179 EVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACL 238

Db        144 EPDAMCTAFQENAEAFMGHYLHEVARRHPYFYGPELLYLADKYTAVLTECCAADDKGACL 203

Qy        239 LPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDL 298

Db        204 TPKLDALKEKALVSAVRQRLKCSSMKKFGERAFKAWAVARMSQTFPNADFAEITKLATDL 263

Qy        299 TKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEM 358

Db        264 TKVTQECCHGDLLECADDRAELAKYMCENQASISSKLQACCDKEMLQKSQCLAEVEHDDM 323

Qy        359 PADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLE 418

Db        324 PADLPALTADFVEDKDVCKNYAEAKDVFLGTFLYEYSRRHPEYSVSLLLRLAKKYEATLE 383

Qy        419 KCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQV 478

Db        384 KCCAEADPHACYGHVFDEFKPLVEEPQNLVKSNCELYEKLGEYGFQNAVLVRYTKKAPQV 443

Qy        479 STPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCT 538

Db        444 STPTLVEAARSLGRVGTHCCALPEKKRLPCVEDYLSAILNRVCLLHEKTPVSEQVTKCCS 503

Qy        539 ESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPK 598

Db        504 GSLVERRPCFSALPVDETYVPKEFKAETFTFHANICTLPEKEKQMEKQTALAELVKHKPQ 563

Qy        599 ATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 643

Db        564 ATEEQLKKVMGDFAEFLEKCCKQEDKEACFSTEGPKLVAESQKAL 608


RESULT 10
A05139
serum albumin - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: A05139; I48638
R;Minghetti, P.P.; Law, S.W.; Dugaiczyk, A.
Mol. Biol. Evol. 2, 347-358, 1985
A;Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudog
A;Reference number: A93055; MUID:88216123; PMID:2452956
A;Accession: A05139
A;Molecule type: mRNA
A;Residues: 1-418 <MIN>
A;Cross-references: UNIPROT:P07724; UNIPARC:UPI000016CBE6; GB:M16111; NID:g191764; PIDN:
R;Boccaccio, C.; Deschatrette, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A;Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the
A;Reference number: I48638; MUID:90269606; PMID:1971802
A;Accession: I48638
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 379-453 <BOC>
A;Cross-references: UNIPARC:UPI000016CEAE; EMBL:X13060; NID:g52939; PIDN:CAA31458.1; PID
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F;123-296/Domain: serum albumin repeat homology <SA2>
F;315-453/Domain: serum albumin repeat homology (fragment) <SA3>

  Query Match             54.5%;  Score 1861;  DB 2;  Length 453;
  Best Local Similarity   72.2%;  Pred. No. 9.1e-115;
  Matches  327;  Conservative   64;  Mismatches   62;  Indels    0;  Gaps    0;

Qy        135 CTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETF 194

Db          1 CAIPNLRENYGELADCCTKQEPERNECFLQHKDDNPSLPPFERPEAEAMCTSFKENPTTF 60

Qy        195 LKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSA 254

Db         61 MGHYLHEVARRHPYFYAPELLYYAEQYNEILTQCCAEADKESCLTPKLDGVKEKALVSSV 120

Qy        255 KQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECA 314

Db        121 RQRMKCSSMQKFGERAFKAWAVARLSQTFPNADFAEITKLATDLTKVNKECCHGDLLECA 180

Qy        315 DDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKD 374

Db        181 DDRAELAKYMCENQATISSKLQTCCDKPLLKKAHCLSEVEHDTMPADLPAIAADFVEDQE 240

Qy        375 VCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVF 434

Db        241 VCKNYAEAKDVFLGTFLYEYSRRHPDYSVSLLLRLAKKYEATLEKCCAEANPPACYGTVL 300

Qy        435 DEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVG 494

Db        301 AEFQPLVEEPKNLVKTNCDLYEKLGEYGFQNAILVRYTQKAPQVSTPTLVEAARNLGRVG 360

Qy        495 SKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVD 554

Db        361 TKCCTLPEDQRLPCVEDYLSAILNRVCLLHEKTPVSEHVTKCCSGSLVERRPCFSALTVD 420

Qy        555 ETYVPKEFNAETFTFHADICTLSEKERQIKKQT 587

Db        421 ETYVPKEFKAETFTFHSDICTLPEKEKQIKKQT 453


RESULT 11
ABCHS
serum albumin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S15571; A05078; A13451
R;Cassady, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A;Reference number: S15571
A;Accession: S15571
A;Molecule type: mRNA
A;Residues: 1-615 <CAS>
A;Cross-references: UNIPROT:P19121; UNIPARC:UPI00001257C1; EMBL:X60688; NID:g63747; PIDN
R;Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A;Title: The 5' noncoding and flanking regions of the avian very low density apolipoprot
A;Reference number: A05078; MUID:83161037; PMID:6187737
A;Accession: A05078
A;Molecule type: DNA
A;Residues: 1-28 <HAC>
A;Cross-references: UNIPARC:UPI000017128E; GB:V00381; NID:g63038; PIDN:CAA23680.1; PID:g
R;Rosen, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A;Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A;Reference number: A13451; MUID:78019943; PMID:911327
A;Accession: A13451
A;Molecule type: protein
A;Residues: 19-23,'M',25-30 <ROS>
A;Cross-references: UNIPARC:UPI000017441B
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thyr
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-26/Domain: propeptide #status predicted <PRO>
F;27-613/Product: serum albumin #status predicted <MAT>
F;32-206/Domain: serum albumin repeat homology <SA1>
F;225-398/Domain: serum albumin repeat homology <SA2>
F;417-596/Domain: serum albumin repeat homology <SA3>
F;30/Binding site: copper (His) #status predicted
F;80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-397,

  Query Match             45.7%;  Score 1562;  DB 1;  Length 615;
  Best Local Similarity   46.9%;  Pred. No. 5.5e-95;
```

```
  Matches  276;  Conservative  118;  Mismatches  192;  Indels    2;  Gaps    2;

Qy         60 RDA-HKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADES 118
              ||| |||| :|||: || || ||| : :| |||||| :| :|   ||| :| : |: |||:|
Db         26 RDAEHKSEIAHRYNDLKEETFKAVAMITFAQYLQRCSYEGLSKLVKDVVDLAQKCVANED 85

Qy        119 AENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPR-LVR 177
              | | | | ::  |::| |  ||::|| ||||| :| :|||||||  |   |: :  |
Db         86 APECSKPLPSIILDEICQVEKLRDSYGAMADCCSKADPERNECFLSFKVSQPDFVQPYQR 145

Qy        178 PEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC 237
              |  ||:|  : ||  :|| ::| :|||||: ||| :| ||  :: |   ||: :|  ||
Db        146 PASDVICQEYQDNRVSFLGHFIYSVARRHPFLYAPAILSFAVDFEHALQSCCKESDVGAC 205

Qy        238 LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTD 297
              |  |   :|:: |  | ||:  |  |::||:| |:|  :  |||::||| |:|||| | |
Db        206 LDTKEIVMREKAKGVSVKQQYFCGILKQFGDRVFQARQLIYLSQKYPKAPFSEVSKFVHD 265

Qy        298 LTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDE 357
                || ||| ||::|| || | :   :|  ||  | |:|:|||||::|:| || | | ||
Db        266 SIGVHKECCEGDMVECMDDMARMMSNLCSQQDVFSGKIKDCCEKPIVERSQCIMEAEFDE 325

Qy        358 MPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTL 417
              |||||||   ::| |:|||::    | |:  |:||||:||||::|: |::|:|| ||: |
Db        326 KPADLPSLVEKYIEDKEVCKSFEAGHDAFMAEFVYEYSRRHPEFSIQLIMRIAKGYESLL 385

Qy        418 EKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQ 477
              ||||   :| ||||   ::    ::| |:::| ||:|    ||  |  ::|:|||||:||
Db        386 EKCCKTDNPAECYANAQEQLNQHIKETQDVVKTNCDLLHDHGEADFLKSILIRYTKKMPQ 445

Qy        478 VSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC 537
              | |  :|  :  :|:|||:  | :|| |:| |||:|::  |   | ||::| |::||
Db        446 VPTDLLLETGKKMTTIGTKCCQLGEDRRMACSEGYLSIVIHDTCRKQETTPINDNVSQCC 505

Qy        538 TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKP 597
              ::   ||||||:|: ||  |||  || : |:|   :|:   :||:: :   |: |:| ||
Db        506 SQLYANRRPCFTAMGVDTKYVPPPFNPDMFSFDEKLCSAPAEERVEGQMKLLINLIKRKP 565

Qy        598 KATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              : |:||:| : | | | |:|||   |  ||| |||   |:  |:| ||:
Db        566 QMTEEQIKTIADGFTAMVDKCCKQSDINTCFGEEGANLIVQSRATLGI 613


RESULT 12
JC4258
alpha-fetoprotein precursor - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
C;Accession: JC4258
R;Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczyk, A.
Gene 162, 213-220, 1995
A;Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to t
A;Reference number: JC4258; MUID:96032345; PMID:7557431
A;Accession: JC4258
A;Molecule type: DNA
A;Residues: 1-609 <NIS>
A;Cross-references: UNIPROT:Q28789; UNIPARC:UPI000012A6FA; GB:U21916; NID:g841311; PIDN:
C;Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
o similar properties and structure.
C;Genetics:
A;Gene: afp
A;Map position: 3p
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-609/Product: alpha-fetoprotein #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;42,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             36.9%;  Score 1260.5;  DB 2;  Length 609;

  Best Local Similarity   39.7%;  Pred. No. 3.4e-75;
  Matches  238;  Conservative  117;  Mismatches  238;  Indels    7;  Gaps    3;

Qy         52 FIAWLVKGRDAHKSE-----VAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEV 106
              |:    : |  |::|     :  ::   | |   |  | |||::|:  :::  |:| :
Db         11 FLLNFTESRTLHRNEYGIASILDSYQCTAEINLTDLATIFFAQFVQEATYKEVSKMVKDA 70

Qy        107 TEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHK 166
                :   || : | ::   | ::||   : | ||  :|||:: |  |: ||| ||
Db         71 LTAIEKPTGDEQSAGCLENQLPAFLEELCREKEILEKYGH-SDCCSQSEEGRHNCFLAHK 129

Qy        167 DDNP-NLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAF 225
                | ::|    ||    | |: :: |||: |::|||||||: ||| :| :| ||
Db        130 KPTPASIPFFQVPEPVTSCEAYEEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKII 189

Qy        226 TECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPK 285
                ||:| :  |   |   :  | : ||   :  || :: || | |:|  | :|||:| |
Db        190 PSCCKAENAVECFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFQAITVTKLSQKFTK 249

Qy        286 AEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLE 345
                | |: ||| |:  ||  || ||:|:|  |   :  |||  ||::|:|: |||:   ||
Db        250 VNFTEIQKLVLDVAHVHEHCCRGDVLDCLQDGEKIMSYICSQQDTLSNKITECCKLTTLE 309

Qy        346 KSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVL 405
              :  ||   |||| |  |     |:  :|  :  :  |::||  |::||:|||| :| :
Db        310 RGQCIIHAENDEKPEGLSPNLNRFLGDRDFNQFSSGEKNIFLASFVHEYSRRHPQLAVSV 369

Qy        406 LLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQN 465
              :||:|| |:  ||||     :| ||  |  :| :  ::|  | |  |::| ||::|||  ||
Db        370 ILRVAKGYQELLEKCFQTENPLECQDKGEEELQKYIQESQALAKRSCGLFQKLGEYYLQN 429

Qy        466 ALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHE 525
              | || |||| ||::  |: ::| :    : ||:  | | : |  :::  ||: ||
Db        430 AFLVAYTKKAPQLTSSELMAITRKMAATAATCCQLSEDKLLACGEGAADIIIGHLCIRHE 489

Qy        526 KTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKK 585
              |||:  | :||| |  |||||||:| |||||||  |: : | || |:|       | |
Db        490 TTPVNPGVGQCCTSSYANRRPCFSSLVVDETYVPPAFSDDKFIFHKDLCQAQGVALQTMK 549

Qy        586 QTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |  |: ||| ||: |:|||:||: ||:  :|||:  ::| ||||||:||:: ::|||||:
Db        550 QEFLINLVKQKPQITEEQLEAVIADFSGLLEKCCQGQEQEVCFAEEGQKLISKTRAALGV 609


RESULT 13
FPHU
alpha-fetoprotein precursor [validated] - human
N;Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C;Species: Homo sapiens (man)
C;Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 09-Jul-2004
C;Accession: A26624; S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042; A0
R;Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczyk, A.
Biochemistry 26, 1332-1343, 1987
A;Title: Structure, polymorphism, and novel repeated DNA elements revealed by a complet
A;Reference number: A26624; MUID:87185438; PMID:2436661
A;Accession: A26624
A;Molecule type: DNA
A;Residues: 1-609 <GIB>
A;Cross-references: UNIPROT:P02771; UNIPARC:UPI00000012A9; GB:M16110; NID:g773678; PIDN
R;McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilghman, S.; Krumlau
Hum. Mol. Genet. 2, 379-384, 1993
A;Title: A G->A substitution in an HNF I binding site in the human alpha-fetoprotein ge
A;Reference number: S37655; MUID:93278385; PMID:7684942
A;Accession: S37655
A;Molecule type: DNA
A;Residues: 1-28 <MCV>
A;Cross-references: UNIPARC:UPI000016A4DF; EMBL:Z19532; NID:g28527; PIDN:CAA79592.1; PI
A;Note: the authors translated the codon TAT for residue 26 as Thr
R;Morinaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A;Title: Primary structures of human alpha-fetoprotein and its mRNA.
A;Reference number: A93961; MUID:83273664; PMID:6192439
A;Accession: A93961
```

A;Molecule type: mRNA
A;Residues: 1-609 <MOR>
A;Cross-references: UNIPARC:UPI00000012A9; GB:J00077; NID:g311348; PIDN:CAA24758.1; PID:
R;Beattie, W.G.; Dugaiczyk, A.
Gene 20, 415-422, 1982
A;Title: Structure and evolution of human alpha-fetoprotein deduced from partial sequenc
A;Reference number: A91497; MUID:83158778; PMID:6187626
A;Accession: A91497
A;Molecule type: mRNA
A;Residues: 429-556 <BEA>
A;Cross-references: UNIPARC:UPI0000174421; GB:J00076
R;Pucci, P.; Siciliano, R.; Malorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Terrana
Biochemistry 30, 5061-5066, 1991
A;Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A;Reference number: A23699; MUID:91242409; PMID:1709810
A;Accession: A23699
A;Molecule type: protein
A;Residues: 19-45;60-97;102-107;122-184;187-249;255-489;507-609 <PUC>
A;Cross-references: UNIPARC:UPI0000174422; UNIPARC:UPI0000174423; UNIPARC:UPI0000174424;
R;Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A;Title: Characterization of in vitro expressed human alpha-fetoprotein as highly reprod
A;Reference number: A61480; MUID:91225826; PMID:1709209
A;Accession: A61480
A;Molecule type: protein
A;Residues: 19-45;63-97;102-107;122-184;187-249;255-489;507-609 <TEC>
A;Cross-references: UNIPARC:UPI0000174422; UNIPARC:UPI0000174424; UNIPARC:UPI0000174425;
R;Yachnin, S.; Hsu, R.; Heinrikson, R.L.; Miller, J.B.
Biochim. Biophys. Acta 493, 418-428, 1977
A;Title: Studies on human alpha-fetoprotein. Isolation and characterization of monomeric
A;Reference number: A90624; MUID:77242506; PMID:70228
A;Accession: A90624
A;Molecule type: protein
A;Residues: 'S',20-22,'S',24-35 <YAC>
A;Cross-references: UNIPARC:UPI000017442A
A;Note: dimeric and trimeric forms have been found in addition to the monomeric form
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 37, 3663-3667, 1977
A;Title: Comparative chemical structure of human alpha-fetoproteins from fetal serum and
A;Reference number: A90757; MUID:78001760; PMID:71198
A;Accession: A90757
A;Molecule type: protein
A;Residues: 'S',20-30,'A',32-37,'A' <AOY>
A;Cross-references: UNIPARC:UPI000017442B
R;Ruoslahti, E.; Pihko, H.; Vaheri, A.; Seppala, M.; Virolainen, M.; Konttinen, A.
Johns Hopkins Med. J. Suppl. 3, 249-255, 1974
A;Title: 20. Alpha fetoprotein: structure and expression in man and inbred mouse strains
A;Reference number: A93042; MUID:75018719; PMID:4138095
A;Accession: A93042
A;Molecule type: protein
A;Residues: 'S',20-24,'Q',26-30,'A',32-35,'E',37-39 <RUO>
A;Cross-references: UNIPARC:UPI000017442C
R;Sakai, M.; Morinaga, T.; Urano, Y.; Watanabe, K.; Wegmann, T.G.; Tamaoki, T.
J. Biol. Chem. 260, 5055-5060, 1985
A;Title: The human alpha-fetoprotein gene. Sequence organization and the 5' flanking reg
A;Reference number: A92520; MUID:85182629; PMID:2580830
A;Contents: annotation; gene, exons and introns
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.
A;Reference number: A90758; MUID:79001617; PMID:80265
A;Contents: annotation; metal binding
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 39, 3571-3574, 1979
A;Title: alpha-Fetoprotein as a carrier protein in plasma and its bilirubin-binding abil
A;Reference number: A90759; MUID:80001710; PMID:89900
A;Contents: annotation; bilirubin binding
C;Comment: AFP is synthesized by the fetal liver and yolk sac. It occurs in the plasma c
o trace amounts after birth. The serum level in adults is usually less than 40 ng/ml. AF
C;Comment: Human AFP binds copper, nickel, and fatty acids as well as, and the bilirubin
properties.
C;Genetics:
A;Gene: GDB:AFP
A;Cross-references: GDB:119660; OMIM:104150
A;Map position: 4q11-4q13
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-609/Product: alpha-fetoprotein #status experimental <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;22/Binding site: copper (His) #status experimental
F;99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
F;249/Binding site: bilirubin (Lys) #status predicted
F;251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 36.8%; Score 1256.5; DB 1; Length 609;
Best Local Similarity 39.5%; Pred. No. 6.2e-75;
Matches 237; Conservative 118; Mismatches 238; Indels 7; Gaps 3;

```
Qy         52 FIAWLVKGRDAHKSE-----VAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEV 106
              |:    : |  |::|     :   ::   | :   |  | |||::|:  :::  |:| :
Db         11 FLLNFTESRTLHRNEYGIASILDSYQCTAEISLADLATIFFAQFVQEATYKEVSKMVKDA 70

Qy        107 TEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHK 166
                :    || :  | ::    | ::||    : | ||  :|||:: |  |: ||| ||
Db         71 LTAIEKPTGDEQSSGCLENQLPAFLEELCHEKEILEKYGH-SDCCSQSEEGRHNCFLAHK 129

Qy        167 DDNP-NLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAF 225
              | ::|    ||    | |: :: |||: |::|||||||: ||| :| :| ||
Db        130 KPTPASIPLFQVPEPVTSCEAYEEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKII 189

Qy        226 TECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPK 285
              ||:| :   |   |   :  | : ||   :  || :: || | |:|  | :|||:| |
Db        190 PSCCKAENAVECFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFQAITVTKLSQKFTK 249

Qy        286 AEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLE 345
              | |: ||| |:  ||  || ||:|:|  |   :  |||  ||::|:|: |||:   ||
Db        250 VNFTEIQKLVLDVAHVHEHCCRGDVLDCLQDGEKIMSYICSQQDTLSNKITECCKLTTLE 309

Qy        346 KSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVL 405
              :  ||   |||| |  |      |: :|  :  :  |::||  |::|||:||||  :| :
Db        310 RGQCIIHAENDEKPEGLSPNLNRFLGDRDFNQFSSGEKNIFLASFVHEYSRRHPQLAVSV 369

Qy        406 LLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQN 465
              :||:|| |:  ||||     :| ||  |  :| :  ::| | | |::| ||::|||  ||
Db        370 ILRVAKGYQELLEKCFQTENPLECQDKGEEELQKYIQESQALAKRSCGLFQKLGEYYLQN 429

Qy        466 ALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHE 525
              | || |||| ||::  |: ::| :     : ||:  | | : | |    :::  ||: ||
Db        430 AFLVAYTKKAPQLTSSELMAITRKMAATAATCCQLSEDKLLACGEGAADIIIGHLCIRHE 489

Qy        526 KTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKK 585
              |||:  | :||| |  ||||||:| |||||||  |: : | || |:|        |  |
Db        490 MTPVNPGVGQCCTSSYANRRPCFSSLVVDETYVPPAFSDDKFIFHKDLCQAQGVALQTMK 549

Qy        586 QTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |  |: ||| ||: |:|||:||: ||:  :||||:  ::| ||||||:||:: ::||||:
Db        550 QEFLINLVKQKPQITEEQLEAVIADFSGLLEKCCQGQEQEVCFAEEGQKLISKTRAALGV 609
```

RESULT 14
FPGO
alpha-fetoprotein precursor - gorilla
C;Species: Gorilla gorilla (gorilla)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A37970
R;Ryan, S.C.; Zielinski, R.; Dugaiczyk, A.
Genomics 9, 60-72, 1991
A;Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primates
A;Reference number: A37970; MUID:91169517; PMID:1706310
A;Accession: A37970
A;Molecule type: DNA
A;Residues: 1-609 <RYA>

A;Cross-references: UNIPROT:P28050; UNIPARC:UPI000012A6F8; GB:M38272; NID:g817963; PIDN:
C;Genetics:
A;Map position: 4q11-12
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-609/Product: alpha-fetoprotein #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;22/Binding site: copper (His) #status predicted
F;99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
F;249/Binding site: bilirubin (Lys) #status predicted
F;251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 36.6%; Score 1249.5; DB 1; Length 609;
Best Local Similarity 39.2%; Pred. No. 1.8e-74;
Matches 235; Conservative 119; Mismatches 239; Indels 7; Gaps 3;

```
Qy         52 FIAWLVKGRDAHKSE-----VAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEV 106
              |:    : |  |::|     :   ::   | :   |  | |||::|:  :::  |:| :
Db         11 FLLNFTESRTLHRNEYGIASILDSYQCTAEISLADLATIFFAQFVQEATYKEVSKMVKDA 70

Qy        107 TEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHK 166
                  :    || :  | ::    | ::||    : | || ::|||:: |  |: ||| ||
Db         71 LTAIEKPTGDEQSAGCLENQLPAFLEELCHEKEILEKYG-LSDCCSQSEEGRHNCFLAHK 129

Qy        167 DDNP-NLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAF 225
               | ::|    ||    | |: :: |||: |::|||||||: ||| :| :| ||
Db        130 KPTPASIPLFQVPEPVTSCEAYEEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKII 189

Qy        226 TECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPK 285
              ||:| :   |   |   :  | : ||   :  || :: || | |:|  | :|||:| |
Db        190 PSCCKAENAVECFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFQAITVTKLSQKFTK 249

Qy        286 AEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLE 345
              | |: ||| |:  ||  || ||:|:|  |   :  |||  ||::|:|: |||:   ||
Db        250 VNFTEIQKLVLDVAHVHEHCCRGDVLDCLQDGEKIMSYICSQQDTLSNKITECCKLTTLE 309

Qy        346 KSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVL 405
              :  ||   |||| |  |      |:  :|  :  :  |::||  |::||:||||  :| :
Db        310 RGQCIIHAENDEKPEGLSPNLNRFLGDRDFNQFSSGEKNIFLASFVHEYSRRHPQLAVSV 369

Qy        406 LLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQN 465
              :||:|| |:  |||| :| ||  |  :| :  ::| | | |::|  ||::|||| ||
Db        370 ILRVAKGYQELLEKCFQTENPLECQDKGEEELQKYIQESQALAKRSCGLFQKLGEYYLQN 429

Qy        466 ALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHE 525
              | || |||| ||:::  |: ::| :    : ||:  | | : | |    :::  ||: ||
Db        430 AFLVAYTKKAPQLTSSELMAITRKMAATAATCCQLSEDKLLACGEGAADIIIGHLCIRHE 489

Qy        526 KTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKK 585
               |||:  | :||| |  |||||||:| |||||||  |: : | || |:|      |  |
Db        490 MTPVNPGVGQCCTSSYANRRPCFSSLVVDETYVPPAFSDDKFIFHKDLCQAQGVALQTMK 549

Qy        586 QTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |  |: ||| ||: |:|||: |: ||:  :||||:  ::| ||||||:||:: :: |||:
Db        550 QEFLINLVKQKPQITEEQLETVIADFSGLLEKCCQGQEQEVCFAEEGQKLISKTRTALGV 609
```

RESULT 15
ABXL72
74K albumin precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: B41682; S02693; A05288
R;Moskaitis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A;Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic ac
e during development.
A;Reference number: A41682; MUID:89313788; PMID:2747653
A;Accession: B41682
A;Molecule type: mRNA
A;Residues: 3-607 <MOS>
A;Cross-references: UNIPROT:P14872; UNIPARC:UPI00001714FB; GB:M21442; NID:g213930; PIDN
R;Schorpp, M.; Doebbeling, U.; Wagner, U.; Ryffel, G.U.
J. Mol. Biol. 199, 83-93, 1988
A;Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Del
A;Reference number: S02692; MUID:88172470; PMID:2451026
A;Accession: S02693
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-48 <SCH>
A;Cross-references: UNIPARC:UPI000017441E; EMBL:Z26826
R;Wolffe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J
Eur. J. Biochem. 146, 489-496, 1985
A;Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilizati
A;Reference number: A05288; MUID:85126974; PMID:3971963
A;Accession: A05288
A;Molecule type: mRNA
A;Residues: 459-502,'L',504-557 <WOL>
A;Cross-references: UNIPARC:UPI000017441F; GB:M28276
A;Note: the authors translated the codon TAT for residue 63 as Thr
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thy
C;Genetics:
A;Introns: 27/1
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: 74K serum albumin #status predicted <MAT>
F;32-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;30/Binding site: copper (His) #status predicted
F;80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392
F;256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.3%; Score 1207.5; DB 1; Length 607;
Best Local Similarity 38.5%; Pred. No. 1e-71;
Matches 232; Conservative 113; Mismatches 243; Indels 15; Gaps 4;

```
Qy         41 SSYLEGQAAKEFIAWLVKGRDA---HKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFE 97
              ||::| :        ::  ||    |  :|  :  | |  || | |   :| ||:|  |
Db         13 SSFIESR--------ILFKRDTDADHHKHIADVYTALTERTFKGLTLAIVSQNLQKCSLE 64

Qy         98 DHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPE 157
              :  |||||: :|||:|: |:: | |:| : ||| ||||    :   |    :|||||:||
Db         65 ELSKLVNEINDFAKSCINDKTPE-CEKPVGTLFFDKLCADPAVGVNYEWSKECCAKQDPE 123

Qy        158 RNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFF 217
              | :||  |:|        ::|| :  |    :: :  |  :::| || ||  | | :|
Db        124 RAQCFKAHRDHEHT---SIKPEPEETCKLLKEHPDDLLSAFIHEEARNHPDLYPPAVLAL 180

Qy        218 AKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVA 277
              |:|      ||:  ||  |   |: :|  :  :   ||   |  |  | |:  || :|
Db        181 TKQYHKLAEHCCEEEDKEKCFSEKMKQLMKQSHSIEDKQHHFCWILDNFPEKVLKALNLA 240

Qy        278 RLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKE 337
              |:| |:|||||         ::|    :||| |: ||  :| :| :: |:::| :||||::
Db        241 RVSHRYPKAEFKLAHNFTEEVTHFIKDCCHDDMFECMTERLELTEHTCQHKDELSSKLEK 300

Qy        338 CCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARR 397
              ||  ||||:::||  :|||::||:|      :| |   ||: |||  :|||| :|:  :|:
Db        301 CCNIPLLERTYCIVTLENDDVPAELSQPITEFTEDPHVCEKYAENNEVFLGRYLHAVSRK 360

Qy        398 HPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQ 457
              | : |   ||: || ||: | |||   :| |||    | |    :|    :||||::  :
Db        361 HQELSEQFLLQSAKEYESLLNKCCKTDNPPECYKDGADRFMNEAKERFAYLKQNCDILHE 420

Qy        458 LGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVL 517
              ||| |:| ||:|||||:|||| ||: ::  :  :|  ||  || :|||||  |::::
Db        421 HGEYLFENELLIRYTKKMPQVSDETLIGIAHQMADIGEHCCAVPENQRMPCAEGDLTILI 480
```

```
Qy        518 NQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLS 577
              ::|   :|| ::: |  |||:|    | ||:||  || |||     :|| |   ||| :
Db        481 GKMCERQKKTFINNHVAHCCTDSYSGMRSCFTALGPDEDYVPPPVTDDTFHFDDKICTAN 540

Qy        578 EKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVA 637
              :||:|  ||  ||:|:|  ||  |  :    :|   |:||| ||: : ||  |   |:
Db        541 DKEKQHIKQKFLVKLIKVSPKLEKNHIDECSAEFLKMVQKCCTADEHQPCFDTEKPVLIE 600

Qy        638 ASQ 640
                |
Db        601 HCQ 603
```

Search completed: April 19, 2006, 12:10:00
Job time : 30.3182 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 11:57:02 ; Search time 170.412 Seconds
(without alignments)
2670.387 Million cell updates/sec

Title: US-10-775-180-447_COPY_30_674
Perfect score: 3417
Sequence: 1 HGEGTFTSDVSSYLEGQAAK.....TCFAEEGKKLVAASQAALGL 645

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3108	91.0	609	1	ALBU_HUMAN	P02768 homo sapien
2	3108	91.0	609	2	Q645G4_HUMAN	Q645g4 homo sapien
3	3108	91.0	609	2	Q5NVH5_PONPY	Q5nvh5 pongo pygma
4	3084	90.3	609	2	Q68DN5_HUMAN	Q68dn5 homo sapien
5	3073	89.9	609	2	Q56G89_HUMAN	Q56g89 homo sapien
6	3066	89.7	627	2	Q5D0D7_HUMAN	Q5d0d7 homo sapien
7	2947	86.2	600	1	ALBU_MACMU	Q28522 macaca mula
8	2627	76.9	608	1	ALBU_FELCA	P49064 felis silve
9	2574	75.3	608	1	ALBU_CANFA	P49822 canis famil
10	2509	73.4	608	2	Q95VB7_SCHMA	Q95vb7 schistosoma
11	2501.5	73.2	607	1	ALBU_EQUAS	Q5xle4 equus asinu
12	2481.5	72.6	607	1	ALBU_HORSE	P35747 equus cabal
13	2469	72.3	608	2	Q5EG48_MICFO	Q5eg48 microtus fo
14	2462	72.1	608	1	ALBU_RABIT	P49065 oryctolagus
15	2460	72.0	608	2	Q5EG49_MICFO	Q5eg49 microtus fo
16	2455.5	71.9	607	1	ALBU_BOVIN	P02769 bos taurus
17	2438	71.3	608	2	Q5U3X3_RAT	Q5u3x3 rattus norv
18	2437.5	71.3	607	1	ALBU_SHEEP	P14639 ovis aries
19	2431	71.1	608	1	ALBU_RAT	P02770 rattus norv
20	2409.5	70.5	607	1	ALBU_PIG	P08835 sus scrofa
21	2392	70.0	608	2	Q6WDN9_CAVPO	Q6wdn9 cavia porce
22	2387.5	69.9	609	1	ALBU_MERUN	O35090 meriones un
23	2383	69.7	608	1	ALBU_MOUSE	P07724 mus musculu
24	2383	69.7	608	2	Q546G4_MOUSE	Q546g4 mus musculu
25	2379.5	69.6	583	2	Q6B3Z0_ELEMA	Q6b3z0 elephas max
26	2379	69.6	608	2	Q8C7H3_MOUSE	Q8c7h3 mus musculu
27	2336	68.4	576	2	Q8C7C7_MOUSE	Q8c7c7 mus musculu
28	1991	58.3	417	2	Q86YG0_HUMAN	Q86yg0 homo sapien
29	1870.5	54.7	396	2	Q8IUK7_HUMAN	Q8iuk7 homo sapien
30	1562	45.7	615	1	ALBU_CHICK	P19121 gallus gall
31	1295.5	37.9	527	2	Q8JIA9_SPHPU	Q8jia9 sphenodon p
32	1260.5	36.9	609	1	FETA_PANTR	Q28789 pan troglod
33	1256.5	36.8	609	1	FETA_HUMAN	P02771 homo sapien
34	1249.5	36.6	609	1	FETA_GORGO	P28050 gorilla gor
35	1242	36.3	609	2	Q8MJU5_CANFA	Q8mju5 canis famil
36	1242	36.3	626	2	Q8UW05_AMBMC	Q8uw05 ambystoma m
37	1218.5	35.7	610	2	Q8MJ76_PIG	Q8mj76 sus scrofa
38	1215.5	35.6	609	2	Q5CZZ1_XENTR	Q5czz1 xenopus tro
39	1207.5	35.3	607	1	ALBU2_XENLA	P14872 xenopus lae
40	1201.5	35.2	607	2	Q642P7_XENLA	Q642p7 xenopus lae
41	1200	35.1	609	1	FETA_HORSE	P49066 equus cabal
42	1183.5	34.6	608	2	Q7TSF3_MARMO	Q7tsf3 marmota mon
43	1178.5	34.5	615	1	FETA_CHICK	P84407 gallus gall
44	1164.5	34.1	606	1	ALBU1_XENLA	P08759 xenopus lae
45	1087	31.8	624	2	Q8UW06_AMBTE	Q8uw06 ambystoma t

ALIGNMENTS

RESULT 1
ALBU_HUMAN
ID ALBU_HUMAN STANDARD; PRT; 609 AA.
AC P02768; O95574; P04277; Q13140; Q6UXK4; Q9P157; Q9P1I7; Q9UHS3;
AC Q9UJZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Serum albumin precursor.
GN Name=ALB;
GN ORFNames=PRO0903, PRO1708, PRO2044, PRO2619, PRO2675, UNQ696/PRO1341;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczyk A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT LYS-420.
RX MEDLINE=82081882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT GLY-121.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczyk A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RT "Human serum albumin.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT HIROSHIMA-1 LYS-378.
RA Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Fetal liver;
RX MEDLINE=21376145; PubMed=11483580; DOI=10.1101/gr.175501;

RA Yu Y., Zhang C., Zhou G., Wu S., Qu X., Wei H., Xing G., Dong C.,
RA Zhai Y., Wan J., Ouyang S., Li L., Zhang S., Zhou K., Zhang Y., Wu C.,
RA He F.;
RT "Gene expression profiling in human fetal liver and identification of
RT tissue- and developmental-stage-specific genes through compiled
RT expression profiles and efficient cloning of full-length cDNAs.";
RL Genome Res. 11:1392-1403(2001).
RN [7]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP PROTEIN SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573; DOI=10.1016/0014-5793(75)80242-0;
RA Meloun B., Moravek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RP PROTEIN SEQUENCE OF 25-609.
RA Brown J.R., Shockley P., Behrens P.Q.;
RL (In) Bing D.H. (eds.);
RL The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RP NUCLEOTIDE SEQUENCE OF 1-455.
RC TISSUE=Liver;
RA Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
RN [11]
RP NUCLEOTIDE SEQUENCE OF 1-26.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 1-167.
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R.L., Watanabe C., Wieand D., Woods K., Xie M.-H.,
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
RN [13]
RP PROTEIN SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075; DOI=10.1016/0014-5793(76)80496-6;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [14]
RP PROTEIN SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994.";
RL Electrophoresis 15:1459-1465(1994).
RN [15]
RP PROTEIN SEQUENCE OF 166-174.
RX MEDLINE=86242180; PubMed=3087352;
RA Mogard M.H., Kobayashi R., Chen C.F., Lee T.D., Reeve J.R. Jr.,
RA Shively J.E., Walsh J.H.;
RT "The amino acid sequence of kinetensin, a novel peptide isolated from
RT pepsin-treated human plasma: homology with human serum albumin,
RT neurotensin and angiotensin.";
RL Biochem. Biophys. Res. Commun. 136:983-988(1986).
RN [16]
RP PROTEIN SEQUENCE OF 166-174.
RX MEDLINE=87194805; PubMed=2437111;
RA Carraway R.E., Mitra S.P., Cochrane D.E.;
RT "Structure of a biologically active neurotensin-related peptide
RT obtained from pepsin-treated albumin(s).";
RL J. Biol. Chem. 262:5968-5973(1987).
RN [17]
RP DISULFIDE BONDS.
RA Saber M.A., Stockbauer P., Moravek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [18]
RP BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin.";
RL Biochem. J. 171:453-459(1978).
RN [19]
RP VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358; DOI=10.1016/0167-4838(87)90088-4;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
RT domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [20]
RP VARIANTS NASKAPI/MERSIN GLU-396 AND MEXICO GLY-574.
RX MEDLINE=87260818; PubMed=3474609;
RA Takahashi N., Takahashi Y., Blumberg B.S., Putnam F.W.;
RT "Amino acid substitutions in genetic variants of human serum albumin
RT and in sequences inferred from molecular cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4413-4417(1987).
RN [21]
RP VARIANTS NAGASAKI-3 GLN-27 YANOMAMA-2 GLU-396; NAGASAKI-2 ASN-399 AND
RP MAKU GLU-565.
RX MEDLINE=88068523; PubMed=3479777;
RA Takshashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
RT Amerindian and Japanese populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [22]
RP VARIANTS FUKUOKA-2 HIS-23; CHRISTCHURCH/HONOLULU-2 GLN-24; TAGLIACOZZO
RP ASN-337 AND ALBUMIN B/OSAKA-2/PHNOM PHEN LYS-594.
RX MEDLINE=89098947; PubMed=2911589;
RA Arai K., Ishioka N., Huss K., Madison J., Putnam F.W.;
RT "Identical structural changes in inherited albumin variants from
RT different populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:434-438(1989).
RN [23]

```
RP   VARIANTS HONOLULU-2 GLN-24; NAGASAKI-1 GLY-293; HIROSHIMA-1 LYS-378;
RP   TOCHIGI LYS-400; HIROSHIMA-2 LYS-406 AND OSAKA-2 LYS-594.

  Query Match             91.0%;  Score 3108;  DB 1;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-188;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 2
Q645G4_HUMAN
ID   Q645G4_HUMAN PRELIMINARY;      PRT;   609 AA.
AC   Q645G4;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Serum albumin.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Yu Z., Fu Y.;
RT   "High Expression HSA in Pichia for Pharmaceutical Use.";
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY728024; AAU21642.1; -; mRNA.
SQ   SEQUENCE   609 AA;  69366 MW;  F88FF61DD242E818 CRC64;

  Query Match             91.0%;  Score 3108;  DB 2;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-188;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 3
Q5NVH5_PONPY
ID   Q5NVH5_PONPY PRELIMINARY;      PRT;   609 AA.
AC   Q5NVH5;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp459F2310.
GN   Name=DKFZp459F2310;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Cortex;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR926060; CAI29688.1; -; mRNA.
DR   SMR; Q5NVH5; 26-608.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
```

```
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Hypothetical protein.
SQ   SEQUENCE   609 AA;  69366 MW;  F88FF61DD242E818 CRC64;

  Query Match             91.0%;  Score 3108;  DB 2;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-188;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 4
Q68DN5_HUMAN
ID   Q68DN5_HUMAN PRELIMINARY;      PRT;   609 AA.
AC   Q68DN5;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein DKFZp779N1935.
GN   Name=DKFZp779N1935;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RG   The German cDNA Consortium;
RA   Bloecker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA   Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR749331; CAH18185.1; -; mRNA.
DR   SMR; Q68DN5; 26-608.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Hypothetical protein.
SQ   SEQUENCE   609 AA;  69402 MW;  3BA3AFF17BF99E94 CRC64;

  Query Match             90.3%;  Score 3084;  DB 2;  Length 609;
  Best Local Similarity   99.1%;  Pred. No. 6.2e-187;
  Matches  581;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              |||:|||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         24 RDAYKSEVAHRFKDLGEENFKALVLIAFAQYPQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHERTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQAALVELVKHRPKA 599
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||:|||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQAALVELVKHRPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 5
Q56G89_HUMAN
ID   Q56G89_HUMAN PRELIMINARY;      PRT;   609 AA.
AC   Q56G89;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Serum albumin.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Li H., Zhang Y., Li X., Yang R., Tang S., Zhang M., Hua S.;
RT   "Homo sapiens serum albumin (HSA) cDNA sequence.";
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY960291; AAX63425.1; -; mRNA.
SQ   SEQUENCE   609 AA;  69084 MW;  39B0ECBB1217A99C CRC64;

  Query Match             89.9%;  Score 3073;  DB 2;  Length 609;
  Best Local Similarity   99.1%;  Pred. No. 3.1e-186;
  Matches  581;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQGLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVGSKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              |||||||||||||||||||||||||||||| ||| ||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDCLSVFLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNGRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 6
Q5D0D7_HUMAN
ID   Q5D0D7_HUMAN PRELIMINARY;      PRT;   627 AA.
AC   Q5D0D7;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   ALB protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC039235; AAH39235.1; -; mRNA.
SQ   SEQUENCE   627 AA;  71704 MW;  271C97408D7EDD04 CRC64;

  Query Match             89.7%;  Score 3066;  DB 2;  Length 627;
  Best Local Similarity   98.5%;  Pred. No. 8.9e-186;
  Matches  577;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||      : | |
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKTCCCKSSCLRL 609


RESULT 7
ALBU_MACMU
ID   ALBU_MACMU     STANDARD;      PRT;   600 AA.
```

```
AC   Q28522;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor (Fragment).
GN   Name=ALB;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93211971; PubMed=8460152;
RA   Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA   Dwulet J., Putnam F.W.;
RT   "cDNA and protein sequence of polymorphic macaque albumins that differ
RT   in bilirubin binding.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; M90463; AAA36906.1; -; mRNA.
DR   PIR; A47391; A47391.
DR   HSSP; P02768; 1E7B.
DR   SMR; Q28522; 19-600.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL       <1     10       By similarity.
FT   PROPEP       11     16       By similarity.
FT   CHAIN        17    600       Serum albumin.
FT   DOMAIN       17    197       Albumin 1.
FT   DOMAIN      204    389       Albumin 2.
FT   DOMAIN      396    587       Albumin 3.
FT   METAL        19     19       Copper (By similarity).
FT   BINDING     256    256       Bilirubin (Potential).
FT   DISULFID     69     78       By similarity.
FT   DISULFID     91    107       By similarity.
FT   DISULFID    106    117       By similarity.
FT   DISULFID    140    185       By similarity.
FT   DISULFID    184    193       By similarity.
FT   DISULFID    216    262       By similarity.
FT   DISULFID    261    269       By similarity.
FT   DISULFID    281    295       By similarity.
FT   DISULFID    294    305       By similarity.
FT   DISULFID    332    377       By similarity.
FT   DISULFID    376    385       By similarity.
FT   DISULFID    408    454       By similarity.
FT   DISULFID    453    464       By similarity.
FT   DISULFID    477    493       By similarity.
FT   DISULFID    492    503       By similarity.
FT   DISULFID    530    575       By similarity.
FT   DISULFID    574    583       By similarity.
FT   NON_TER       1      1
SQ   SEQUENCE   600 AA;  67881 MW;  E45C871A670E740B CRC64;

  Query Match             86.2%;  Score 2947;  DB 1;  Length 600;
  Best Local Similarity   93.5%;  Pred. No. 2.9e-178;
  Matches  546;  Conservative   23;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              || |||||||||||||||:|| |||:||:|||||||||:|||||||||||||||||||||
Db         16 RDTHKSEVAHRFKDLGEEHFKGLVLVAFSQYLQQCPFEEHVKLVNEVTEFAKTCVADESA 75

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db         76 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPPLVRPE 135

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              |||||||||||| |||||||||:|||||||||||||||| |||||| ||||||||||||
Db        136 VDVMCTAFHDNEATFLKKYLYEVARRHPYFYAPELLFFAARYKAAFAECCQAADKAACLL 195

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              |||||||||||||||||||||||||||:|||||||||||||:|||||||||||||||||
Db        196 PKLDELRDEGKASSAKQRLKCASLQKFGDRAFKAWAVARLSQKFPKAEFAEVSKLVTDLT 255

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              |||||||||||||||||||||||:|||||||||||||:|||||||||:|||||||||
Db        256 KVHTECCHGDLLECADDRADLAKYMCENQDSISSKLKECCDKPLLEKSHCLAEVENDEMP 315

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              |||||||||:|||||||||||||||||||||||||||||||||||:|||||| || ||||
Db        316 ADLPSLAADYVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVMLLLRLAKAYEATLEK 375

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              |||||||||||||||||:|||||||||:|||||||||||||||||||||||||||||||
Db        376 CCAAADPHECYAKVFDEFQPLVEEPQNLVKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 435

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              |||||||||||||:|||| |||||||||||||||:||||||||::|||||||
Db        436 TPTLVEVSRNLGKVGAKCCKLPEAKRMPCAEDYLSVVLNRLCVLHEKTPVSEKVTKCCTE 495

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              |||||||||||||:|| |||| |||||||||||:|||||||:|:|||||||||||||||
Db        496 SLVNRRPCFSALELDEAYVPKAFNAETFTFHADMCTLSEKEKQVKKQTALVELVKHKPKA 555

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 643
              |||||| |||:|||||||||||||| |||||| | ||||||||
Db        556 TKEQLKGVMDNFAAFVEKCCKADDKEACFAEEGPKFVAASQAAL 599


RESULT 8
ALBU_FELCA
ID   ALBU_FELCA     STANDARD;      PRT;   608 AA.
AC   P49064; Q7YSG3;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor (Allergen Fel d 2).
GN   Name=ALB;
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
OC   Felinae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=96194824; PubMed=8647469; DOI=10.1016/0378-1119(95)00851-9;
RA   Hilger C., Grigioni F., Kohnen M., Hentges F.;
RT   "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL   Gene 169:295-296(1996).
RN   [2]
RP   NUCLEOTIDE SEQUENCE OF 25-608.
RC   TISSUE=Liver;
RA   Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA   Rumpold H., Valenta R., Spitzauer S.;
RT   "Escherichia coli expression and purification of recombinant cat
```

```
RT   albumin:IgE recognition, induction of basophil activation and
RT   lymphoproliferative responses in atopic patients.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; X84842; CAA59279.1; -; mRNA.
DR   EMBL; AJ487677; CAD32275.1; -; mRNA.
DR   PIR; JC4660; S57632.
DR   HSSP; P02768; 1E7B.
DR   SMR; P49064; 26-608.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL       1    18       By similarity.
FT   PROPEP      19    24       By similarity.
FT   CHAIN       25   608       Serum albumin.
FT   DOMAIN      25   205       Albumin 1.
FT   DOMAIN     212   397       Albumin 2.
FT   DOMAIN     404   595       Albumin 3.
FT   METAL       27    27       Copper.
FT   DISULFID    77    86       By similarity.
FT   DISULFID    99   115       By similarity.
FT   DISULFID   114   125       By similarity.
FT   DISULFID   148   193       By similarity.
FT   DISULFID   192   201       By similarity.
FT   DISULFID   224   270       By similarity.
FT   DISULFID   269   277       By similarity.
FT   DISULFID   289   303       By similarity.
FT   DISULFID   302   313       By similarity.
FT   DISULFID   340   385       By similarity.
FT   DISULFID   384   393       By similarity.
FT   DISULFID   416   462       By similarity.
FT   DISULFID   461   472       By similarity.
FT   DISULFID   485   501       By similarity.
FT   DISULFID   500   511       By similarity.
FT   DISULFID   538   583       By similarity.
FT   DISULFID   582   591       By similarity.
FT   CONFLICT    75    75       K -> N (in Ref. 2).
FT   CONFLICT    94    94       L -> F (in Ref. 2).
FT   CONFLICT   186   186       K -> R (in Ref. 2).
FT   CONFLICT   251   251       E -> D (in Ref. 2).
FT   CONFLICT   282   282       A -> E (in Ref. 2).
FT   CONFLICT   331   331       V -> A (in Ref. 2).
SQ   SEQUENCE   608 AA;  68659 MW;  07E629CAC5F60E5F CRC64;

  Query Match             76.9%;  Score 2627;  DB 1;  Length 608;
  Best Local Similarity   80.1%;  Pred. No. 5.4e-158;
  Matches  483;  Conservative   53;  Mismatches   57;  Indels   10;  Gaps    1;

Qy         41 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 100
              |:|  |           |  |:||:||:|||| |||||:|: |||:||:|||||||||||
Db         15 SAYSRG----------VTRREAHQSEIAHRFNDLGEEHFRGLVLVAFSQYLQQCPFEDHV 64

Qy        101 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 160
              |||||||||| |||||:|| ||:|||| | ||||||||:||: |||||||| |:||||||
Db         65 KLVNEVTEFAKGCVADQSAANCEKSLHELLGDKLCTVASLRDKYGEMADCCEKKEPERNE 124

Qy        161 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 220
              |||||||||    :|| || | ||||||:||: || ||||||||||||||||||::|:
Db        125 CFLQHKDDNPGFGQLVTPEADAMCTAFHENEQRFLGKYLYEIARRHPYFYAPELLYYAEE 184

Qy        221 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 280
              ||  |||||:|||||||| ||:| ||::  |||||:||||||||||||||||:|||||
Db        185 YKGVFTECCEAADKAACLTPKVDALREKVLASSAKERLKCASLQKFGERAFKAWSVARLS 244

Qy        281 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 340
              |:||||||:|||||| |:| ||||||||||||||||||||||||||:||||||
Db        245 QKFPKAEFAEISKLVTDLAKIHKECCHGDLLECADDRADLAKYICENQDSISTKLKECCG 304

Qy        341 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 400
              ||:|||||||:||| ||:||||| || |||| |:||||| |||||||| |||||:||||:
Db        305 KPVLEKSHCISEVERDELPADLPPLAVDFVEDKEVCKNYQEAKDVFLGTFLYEYSRRHPE 364

Qy        401 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 460
              ||| ||||||| || ||||||| ||  ||| ||||||||||| ||:| ||||||:|||
Db        365 YSVSLLLRLAKEYEATLEKCCATDDPPACYAHVFDEFKPLVEEPHNLVKTNCELFEKLGE 424

Qy        461 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 520
              | |||||||||||||||||||||:|||||||| ||||:|: ||||||||||:|
Db        425 YGFQNALLVRYTKKVPQVSTPTLVEVSRSLGKVGSKCCTHPEAERLSCAEDYLSVVLNRL 484

Qy        521 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 580
              |||||||||:|||||||||||||||||:||||||||:|||||||||:||| | |
Db        485 CVLHEKTPVSERVTKCCTESLVNRRPCFSALQVDETYVPKEFSAETFTFHADLCTLPEAE 544

Qy        581 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 640
              :|||||:|||||:|||||||:|||| || || :||:||| |:||| |||||| |||||:|
Db        545 KQIKKQSALVELLKHKPKATEEQLKTVMGDFGSFVDKCCAAEDKEACFAEEGPKLVAAAQ 604

Qy        641 AAL 643
              |||
Db        605 AAL 607


RESULT 9
ALBU_CANFA
ID   ALBU_CANFA     STANDARD;      PRT;   608 AA.
AC   P49822; O77705; Q9TSZ4;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor (Allergen Can f 3).
GN   Name=ALB;
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC   Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA].
RC   STRAIN=Beagle; TISSUE=Liver;
RA   Hilger C.;
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=20148667; PubMed=10669848; DOI=10.1016/S0091-6749(00)90077-0;
RA   Pandjaitan B., Swoboda I., Brandejsky-Pichler F., Rumpold H.,
RA   Valenta R., Spitzauer S.;
RT   "Escherichia coli expression and purification of recombinant dog
RT   albumin, a cross-reactive animal allergen.";
RL   J. Allergy Clin. Immunol. 105:279-285(2000).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [MRNA].
RC   STRAIN=Beagle; TISSUE=Liver;
RA   Miyake M., Okazaki M., Iwabuchi S.;
```

```
RT   "Isolation of a cDNA encoding canine serum albumin.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   PROTEIN SEQUENCE OF 25-48.
RX   MEDLINE=75011422; PubMed=4414013;
RA   Dixon J.W., Sarkar B.;
RT   "Isolation, amino acid sequence and copper(II)-binding properties of
RT   peptide (1-24) of dog serum albumin.";
RL   J. Biol. Chem. 249:5872-5877(1974).
RN   [5]
RP   PROTEIN SEQUENCE OF 25-38.
RC   TISSUE=Heart;
RX   MEDLINE=98163340; PubMed=9504812;
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
RN   [6]
RP   NUCLEOTIDE SEQUENCE OF 215-478.
RC   TISSUE=Salivary gland;
RX   MEDLINE=94201492; PubMed=7512102;
RA   Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,
RA   Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
RT   "Molecular characterization of dog albumin as a cross-reactive
RT   allergen.";
RL   J. Allergy Clin. Immunol. 93:614-627(1994).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AJ133489; CAB64867.1; -; mRNA.
DR   EMBL; Y17737; CAA76841.1; -; mRNA.
DR   EMBL; AB090854; BAC10663.1; -; mRNA.
DR   EMBL; S72946; AAB30434.1; -; mRNA.
DR   HSSP; P02768; 1E7B.
DR   SMR; P49822; 26-607.
DR   HSC-2DPAGE; P49822; DOG.
DR   Ensembl; ENSCAFG00000003016; Canis familiaris.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Allergen; Copper; Direct protein sequencing; Lipid-binding;
KW   Metal-binding; Repeat; Signal.
FT   SIGNAL       1     18       Potential.
FT   PROPEP      19     24
FT   CHAIN       25    608       Serum albumin.
FT   DOMAIN      25    205       Albumin 1.
FT   DOMAIN     212    397       Albumin 2.
FT   DOMAIN     404    595       Albumin 3.
FT   METAL       27     27       Copper (By similarity).
FT   DISULFID    77     86       By similarity.
FT   DISULFID    99    115       By similarity.
FT   DISULFID   114    125       By similarity.
FT   DISULFID   148    193       By similarity.
FT   DISULFID   192    201       By similarity.
FT   DISULFID   224    270       By similarity.
FT   DISULFID   269    277       By similarity.
FT   DISULFID   289    303       By similarity.
FT   DISULFID   302    313       By similarity.
FT   DISULFID   340    385       By similarity.
FT   DISULFID   384    393       By similarity.
FT   DISULFID   416    462       By similarity.
FT   DISULFID   461    472       By similarity.
FT   DISULFID   485    501       By similarity.
FT   DISULFID   500    511       By similarity.
FT   DISULFID   538    583       By similarity.
FT   DISULFID   582    591       By similarity.
FT   CONFLICT     1     26       MKWVTFISLFFLFSSAYSRGLVRREA -> MDT (in
FT                               Ref. 2).
FT   CONFLICT   146    146       A -> R (in Ref. 2).
FT   CONFLICT   206    206       I -> T (in Ref. 2).
FT   CONFLICT   349    349       V -> A (in Ref. 2).
FT   CONFLICT   359    359       A -> S (in Ref. 2 and 6).
FT   CONFLICT   448    448       V -> VV (in Ref. 6).
FT   CONFLICT   474    474       E -> D (in Ref. 2 and 6).
SQ   SEQUENCE   608 AA;  68605 MW;  3DB012FF7C979CF3 CRC64;

  Query Match             75.3%;  Score 2574;  DB 1;  Length 608;
  Best Local Similarity   78.4%;  Pred. No. 1.2e-154;
  Matches  473;  Conservative   57;  Mismatches   63;  Indels   10;  Gaps    2;

Qy         41 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 100
              |:|  |          ||: |:|:|||:|||: |||||:|: |||:||:||||||||||||
Db         15 SAYSRG---------LVR-REAYKSEIAHRYNDLGEEHFRGLVLVAFSQYLQQCPFEDHV 64

Qy        101 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 160
              ||  ||||||| | |:|| ||||||||||||||||||:||: ||:||||| ||||:|||
Db         65 KLAKEVTEFAKACAAEESGANCDKSLHTLFGDKLCTVASLRDKYGDMADCCEKQEPDRNE 124

Qy        161 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 220
              ||| ||||||  | || || | :| || |||: || ||||||||||||||||||::|::
Db        125 CFLAHKDDNPGFPPLVAPEPDALCAAFQDNEQLFLGKYLYEIARRHPYFYAPELLYYAQQ 184

Qy        221 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 280
              ||  | ||||||||||| ||:: ||::   ||||:| |||||||||:||||||:|||||
Db        185 YKGVFAECCQAADKAACLGPKIEALREKVLLSSAKERFKCASLQKFGDRAFKAWSVARLS 244

Qy        281 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 340
              ||||||:|||:||:|||||||| ||||||||||||||||||:||||||||:||||||:
Db        245 QRFPKADFAEISKVVTDLTKVHKECCHGDLLECADDRADLAKYMCENQDSISTKLKECCD 304

Qy        341 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 400
              ||:|||| |:|||| ||:| |||||||||||| |:||||| |||||||| ||||||||:
Db        305 KPVLEKSQCLAEVERDELPGDLPSLAADFVEDKEVCKNYQEAKDVFLGTFLYEYARRHPE 364

Qy        401 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 460
              ||| ||||||| || |||||||  ||  ||||| |||||||:|||||:| ||||||:|||
Db        365 YSVSLLLRLAKEYEATLEKCCATDDPPTCYAKVLDEFKPLVDEPQNLVKTNCELFEKLGE 424

Qy        461 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 520
              | |||||||||||| |||||||||||| |||||:|||| ||::|| ||||:|||||:|
Db        425 YGFQNALLVRYTKKAPQVSTPTLVEVSRKLGKVGTKCCKKPESERMSCAEDFLSVVLNRL 484

Qy        521 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 580
              ||||||||||:||||||:|||||||||||| |||||||||||||||||||||:||| | |
Db        485 CVLHEKTPVSERVTKCCSESLVNRRPCFSGLEVDETYVPKEFNAETFTFHADLCTLPEAE 544

Qy        581 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 640
              :|:|||||||||:|||||| |||| || || ||||||| |::|| ||:||| |||||:|
Db        545 KQVKKQTALVELLKHKPKATDEQLKTVMGDFGAFVEKCCAAENKEGCFSEEGPKLVAAAQ 604

Qy        641 AAL 643
              |||
Db        605 AAL 607


RESULT 10
Q95VB7_SCHMA
ID   Q95VB7_SCHMA PRELIMINARY;      PRT;   608 AA.
AC   Q95VB7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Albumin.
OS   Schistosoma mansoni (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6183;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF418550; AAL08579.1; -; mRNA.
DR   HSSP; P02768; 1HK1.
DR   SMR; Q95VB7; 26-608.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0008289; F:lipid binding; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 2.
SQ   SEQUENCE   608 AA;  68225 MW;  E5EABB28E1C66E54 CRC64;

  Query Match             73.4%;  Score 2509;  DB 2;  Length 608;
  Best Local Similarity   76.4%;  Pred. No. 1.6e-150;
  Matches  446;  Conservative   79;  Mismatches   59;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              |||||||:|||||||||::|| |||||:|:||:||:|:|||||||||:||||||||||
Db         24 RDAHKSEIAHRFKDLGEQHFKGLVLIAFSQFLQKCPYEEHVKLVNEVTDFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              |||||||||||||||| : |||::|||:|||||:|||||||||:||||:||||  |||:
Db         84 ENCDKSLHTLFGDKLCAIPTLRDSYGELADCCAKKEPERNECFLKHKDDHPNLPPFVRPD 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
               : |||:| :|  ||:  ||:|:|||||||||||::|::| |  ||||  ||||||:
Db        144 AEAMCTSFQENAVTFMGHYLHEVARRHPYFYAPELLYYAEKYSAIMTECCGEADKAACIT 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              |||| |:::  |||  |||||:|||:||:||||||||||:||:||||:|||::|| ||||
Db        204 PKLDALKEKALASSVNQRLKCSSLQRFGQRAFKAWAVARMSQKFPKADFAEITKLATDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              |:  |||||||||||||||:||||:|||| |||||: ||:||:|:||||::|||||::|
Db        264 KLTEECCHGDLLECADDRAELAKYMCENQASISSKLQACCDKPVLKKSHCLSEVENDDLP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              |||||||||||| |:||||||||||||| |||||||||||||| |||||| || ||||
Db        324 ADLPSLAADFVEDKEVCKNYAEAKDVFLGTFLYEYARRHPDYSVALLLRLAKKYEATLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||| |||  || || |||:|||||:||:| ||||||:|||| |||||:||||:| ||||
Db        384 CCAEADPSACYGKVLDEFQPLVEEPKNLVKANCELFEKLGEYGFQNALIVRYTQKAPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              |||||| :|||||||||||  |||:|:|| |||:| :||::||||||||::||||||
Db        444 TPTLVEAARNLGKVGSKCCVLPEAQRLPCVEDYISAILNRVCVLHEKTPVSEQVTKCCTG 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              |:| |||||||| |||||||||| ||||||||||:| |||:|:||| ||||||||||||
Db        504 SVVERRPCFSALPVDETYVPKEFKAETFTFHADICSLPEKEKQMKKQAALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 643
              |  ||: |: :| ||::|||||:||| ||:|:| ||||:|||||
Db        564 TGPQLRTVLGEFTAFLDKCCKAEDKEACFSEDGPKLVASSQAAL 607

RESULT 11
```

```
ALBU_EQUAS
ID   ALBU_EQUAS     STANDARD;      PRT;   607 AA.
AC   Q5XLE4;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor.
GN   Name=ALB;
OS   Equus asinus (Donkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9793;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Li H., Tang Y., Pingfan R.;
RT   "Full-length cDNA sequence of serum albumin of donkey (Equus asinus)
RT   and its structure analysis.";
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AY754333; AAV28861.1; -; mRNA.
DR   SMR; Q5XLE4; 27-607.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL        1     18       By similarity.
FT   PROPEP       19     24       By similarity.
FT   CHAIN        25    607       Serum albumin.
FT   DOMAIN       25    204       Albumin 1.
FT   DOMAIN      211    396       Albumin 2.
FT   DOMAIN      403    594       Albumin 3.
FT   METAL        27     27       Copper (By similarity).
FT   DISULFID     77     86       By similarity.
FT   DISULFID     99    115       By similarity.
FT   DISULFID    114    125       By similarity.
FT   DISULFID    147    192       By similarity.
FT   DISULFID    191    200       By similarity.
FT   DISULFID    223    269       By similarity.
FT   DISULFID    268    276       By similarity.
FT   DISULFID    288    302       By similarity.
FT   DISULFID    301    312       By similarity.
FT   DISULFID    339    384       By similarity.
FT   DISULFID    383    392       By similarity.
FT   DISULFID    415    461       By similarity.
FT   DISULFID    460    471       By similarity.
FT   DISULFID    484    500       By similarity.
FT   DISULFID    499    510       By similarity.
FT   DISULFID    537    582       By similarity.
FT   DISULFID    581    590       By similarity.
SQ   SEQUENCE   607 AA;  68539 MW;  7099E1E08E3C426A CRC64;

  Query Match             73.2%;  Score 2501.5;  DB 1;  Length 607;
  Best Local Similarity   75.0%;  Pred. No. 4.8e-150;
  Matches  452;  Conservative   71;  Mismatches   69;  Indels   11;  Gaps    2;
```

```
Qy         41 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 100
              |:|  |           |  || ||||:|||| ||||::|| |||:||:||||||||||||
Db         15 SAYFRG----------VLRRDTHKSEIAHRFNDLGEKHFKGLVLVAFSQYLQQCPFEDHV 64

Qy        101 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 160
              |||||||||||  | ||||||||||||||||||||||||||||| ||||:|||| ||||||||
Db         65 KLVNEVTEFAKKCAADESAENCDKSLHTLFGDKLCTVATLRATYGELADCCEKQEPERNE 124

Qy        161 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 220
              ||| ||||:||||:| :|| |  | || :: : || |||||:|||||||| ||||| |:
Db        125 CFLTHKDDHPNLPKL-KPEPDAQCAAFQEDPDKFLGKYLYEVARRHPYFYGPELLFHAEE 183

Qy        221 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 280
              ||| ||||| | ||| ||:|||| |::    ||||:||||:| ||||||||||:|||||
Db        184 YKADFTECCPADDKAGCLIPKLDALKERILLSSAKERLKCSSFQKFGERAFKAWSVARLS 243

Qy        281 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 340
              |:||||:||||||:|||||||| ||||||||||||||||| |||||:||||| ||| ||:
Db        244 QKFPKADFAEVSKIVTDLTKVHKECCHGDLLECADDRADLTKYICEHQDSISGKLKACCD 303

Qy        341 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 400
              ||||:|||||||: |::|:|||:||||| | |::||:| :|||||| |||||:|||||
Db        304 KPLLQKSHCIAEVKEDDLPSDLPALAADFAEDKEICKHYKDAKDVFLGTFLYEYSRRHPD 363

Qy        401 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 460
              ||| ||||:||||| ||||||| |||  ||| |||:| ||||||::|:|:|:|||::||
Db        364 YSVSLLLRIAKTYEATLEKCCAEADPPACYATVFDQFTPLVEEPKSLVKKNCDLFEEVGE 423

Qy        461 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 520
              | |||||:|||||| |||||||||:| | ||||||:||| ||::|:||:|::|:: ||:|
Db        424 YDFQNALIVRYTKKAPQVSTPTLVEIGRTLGKVGSRCCKLPESERLPCSENHLALALNRL 483

Qy        521 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 580
              ||||||||||:::|||||:||  |||||||||:|| |:|||| ||||||||||||| | |
Db        484 CVLHEKTPVSEKITKCCTDSLAERRPCFSALELDEGYIPKEFKAETFTFHADICTLPEDE 543

Qy        581 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 640
              :|||||:|| |||||||||||||| |: :|:||| ||| |:||| |||||| ||||:||
Db        544 KQIKKQSALAELVKHKPKATKEQLKTVLGNFSAFVAKCCGAEDKEACFAEEGPKLVASSQ 603

Qy        641 AAL 643
               ||
Db        604 LAL 606
```

```
RESULT 12
ALBU_HORSE
ID   ALBU_HORSE     STANDARD;      PRT;   607 AA.
AC   P35747;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor (Allergen Equ c 3).
GN   Name=ALB;
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC   TISSUE=Liver;
RX   MEDLINE=93345495; PubMed=8344282;
RA   Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT   "X-ray and primary structure of horse serum albumin (Equus caballus)
RT   at 0.27-nm resolution.";
RL   Eur. J. Biochem. 215:205-212(1993).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human. Binds to IgE.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; X74045; CAA52194.1; -; mRNA.
DR   PIR; S34053; ABHOS.
DR   HSSP; P02768; 1HK1.
DR   SMR; P35747; 27-607.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL       1     18       By similarity.
FT   PROPEP      19     24       By similarity.
FT   CHAIN       25    607       Serum albumin.
FT   DOMAIN      25    204       Albumin 1.
FT   DOMAIN     211    396       Albumin 2.
FT   DOMAIN     403    594       Albumin 3.
FT   METAL       27     27       Copper (By similarity).
FT   DISULFID    77     86
FT   DISULFID    99    115
FT   DISULFID   114    125
FT   DISULFID   147    192
FT   DISULFID   191    200
FT   DISULFID   223    269
FT   DISULFID   268    276
FT   DISULFID   288    302
FT   DISULFID   301    312
FT   DISULFID   339    384
FT   DISULFID   383    392
FT   DISULFID   415    461
FT   DISULFID   460    471
FT   DISULFID   484    500
FT   DISULFID   499    510
FT   DISULFID   537    582
FT   DISULFID   581    590
SQ   SEQUENCE   607 AA;  68599 MW;  256F6E830A1B90C5 CRC64;
```

```
  Query Match             72.6%;  Score 2481.5;  DB 1;  Length 607;
  Best Local Similarity   74.6%;  Pred. No. 8.8e-149;
  Matches  450;  Conservative   70;  Mismatches   72;  Indels   11;  Gaps    2;

Qy         41 SSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHV 100
              |:|  |           |  || ||||:|||| ||||::|| |||:||:||||||||||||
Db         15 SAYSRG----------VLRRDTHKSEIAHRFNDLGEKHFKGLVLVAFSQYLQQCPFEDHV 64

Qy        101 KLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE 160
              |||||||||||  | ||||||||||||||||||||||||||||| ||||:|||| ||||||||
Db         65 KLVNEVTEFAKKCAADESAENCDKSLHTLFGDKLCTVATLRATYGELADCCEKQEPERNE 124

Qy        161 CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKR 220
              ||| ||||:||||:| :|| |  | || :: : || |||||:|||||||| ||||| |:
Db        125 CFLTHKDDHPNLPKL-KPEPDAQCAAFQEDPDKFLGKYLYEVARRHPYFYGPELLFHAEE 183

Qy        221 YKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS 280
              ||| ||||| | || ||:|||| |::    ||||:||||:| | ||||| |||:|||||
Db        184 YKADFTECCPADDKLACLIPKLDALKERILLSSAKERLKCSSFQNFGERAVKAWSVARLS 243

Qy        281 QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE 340
              |:||||:||||||:|||||||| ||||||||||||||||||||:||||| ||| ||:
Db        244 QKFPKADFAEVSKIVTDLTKVHKECCHGDLLECADDRADLAKYICEHQDSISGKLKACCD 303

Qy        341 KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPD 400
              ||||:|||||||: |::|:|||:||||| | |::||:| :|||||| |||||:|||||
Db        304 KPLLQKSHCIAEVKEDDLPSDLPALAADFAEDKEICKHYKDAKDVFLGTFLYEYSRRHPD 363
```

```
Qy        401 YSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE 460
              ||| ||||:||||| ||||||| |||  ||  |||:| ||||||::|:|:||:|||::||
Db        364 YSVSLLLRIAKTYEATLEKCCAEADPPACYRTVFDQFTPLVEEPKSLVKKNCDLFEEVGE 423

Qy        461 YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQL 520
              | |||||:|||||| |||||||||:| | ||||||:||| ||::|:||:|::|:: ||:|
Db        424 YDFQNALIVRYTKKAPQVSTPTLVEIGRTLGKVGSRCCKLPESERLPCSENHLALALNRL 483

Qy        521 CVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE 580
              |||||||||:::|||||:||  |||||||||:|| |||||| |||||||||||||| | |
Db        484 CVLHEKTPVSEKITKCCTDSLAERRPCFSALELDEGYVPKEFKAETFTFHADICTLPEDE 543

Qy        581 RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ 640
              :|||||:|| |||||||||||||| |: :|:||| |||  :||| |||||| ||||:||
Db        544 KQIKKQSALAELVKHKPKATKEQLKTVLGNFSAFVAKCCGREDKEACFAEEGPKLVASSQ 603

Qy        641 AAL 643
              ||
Db        604 LAL 606


RESULT 13
Q5EG48_MICFO
ID   Q5EG48_MICFO PRELIMINARY;      PRT;   608 AA.
AC   Q5EG48;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Albumin.
OS   Microtus fortis calamorum.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Arvicolinae; Microtus.
OX   NCBI_TaxID=311220;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Hu W.-X., Wu G.-J., Qin Z.-Q., Luo S.-Q.;
RT   "Albumin gene of Microtus Fortis calamorum liver.";
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY885265; AAW79113.1; -; mRNA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 2.
SQ   SEQUENCE   608 AA;  68308 MW;  B04A061339494D3F CRC64;

  Query Match           72.3%;  Score 2469;  DB 2;  Length 608;
  Best Local Similarity   76.2%;  Pred. No. 5.5e-148;
  Matches  445;  Conservative   65;  Mismatches   74;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              |||||||:|||: |||| || |||| |||:||:||:|:|:|||||||:||| | |||||
Db         24 RDAHKSEIAHRYNDLGEKYFKGLVLITFAQHLQKCPYEEHIKLVNEVTDFAKACAADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              |||||||||||||| :  | : | |||:||||||||||:|||:|||| ||||  ||||
Db         84 ENCDKSLHTLFGDKLCAIPNLGDNYAEMAECCAKQEPERNQCFLKHKDDKPNLPPFVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              :||||:| :|  :|:  ||: :||||||||||||::|::  |  ||||  |||||||
Db        144 AEVMCTSFQENPASFMGHYLHAVARRHPYFYAPELLYYAEKQSAIMTECCAEADKAACLG 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              |||| |:::   || ||||||:||:||||||||||||||:||:||||:|||::|| ||||
Db        204 PKLDALKEKALLSSVKQRLKCSSLEKFGERAFKAWAVARMSQKFPKADFAEITKLATDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              || |||||||||||||| :||||:|:|| :|||||  ||:||:|:|:||:|||::||||
Db        264 KVTQECCHGDLLECADDRLELAKYMCDNQATISSKLHTCCDKPVLQKAHCLAEVDHDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              |||  || |||| |||||||||||||||| ||||||||||||| ||||||| || ||||
Db        324 ADLTPLADDFVEDKDVCKNYAEAKDVFLGTFLYEYARRHPDYSVALLLRLAKKYEATLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||| |||| || |||||||:||||||:||:| |||||||:|||| ||||||||||:| ||||
Db        384 CCAEADPHACYGKVFDEFQPLVEEPKNLVKANCELFEKLGEYGFQNALLVRYTQKAPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              |||||||||||||:|||  ||| |:|| ||||| :|| |||||||||||::||||||
Db        444 TPTLVEVSRNLGKVGTKCCALPEADRLPCVEDYLSAILNHLCVLHEKTPVSEQVTKCCTG 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              |:| |||||||| ||| |||||| ||:|||||||||| :||:|||||||| |||||||||
Db        504 SVVERRPCFSALPVDEHYVPKEFKAESFTFHADICTLPDKEKQIKKQTALAELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 643
              | :||| || :|:||:|||||||||| ||:||| |||| |||||
Db        564 TGDQLKTVMGEFSAFLEKCCKADDKEACFSEEGPKLVATSQAAL 607


RESULT 14
ALBU_RABIT
ID   ALBU_RABIT     STANDARD;      PRT;   608 AA.
AC   P49065;
DT   01-FEB-1996 (Rel. 33, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serum albumin precursor.
GN   Name=ALB;
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC   Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=New Zealand white; TISSUE=Liver;
RX   MEDLINE=97275135; PubMed=9129029;
RA   Syed S., Schuyler P.D., Kulczycky M., Sheffield W.P.;
RT   "Potent antithrombin activity and delayed clearance from the
RT   circulation characterize recombinant hirudin genetically fused to
RT   albumin.";
RL   Blood 89:3243-3252(1997).
RN   [2]
RP   SEQUENCE REVISION TO 322-323 AND 506-507.
RA   Sheffield W.P.;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; U18344; AAB58347.2; -; mRNA.
DR   HSSP; P02768; 1E7B.
DR   SMR; P49065; 26-608.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
```

```
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL       1     18       By similarity.
FT   PROPEP      19     24       By similarity.
FT   CHAIN       25    608       Serum albumin.
FT   DOMAIN      25    205       Albumin 1.
FT   DOMAIN     212    397       Albumin 2.
FT   DOMAIN     404    595       Albumin 3.
FT   METAL       27     27       Copper.
FT   DISULFID    77     86       By similarity.
FT   DISULFID    99    115       By similarity.
FT   DISULFID   114    125       By similarity.
FT   DISULFID   148    193       By similarity.
FT   DISULFID   192    201       By similarity.
FT   DISULFID   224    270       By similarity.
FT   DISULFID   269    277       By similarity.
FT   DISULFID   289    303       By similarity.
FT   DISULFID   302    313       By similarity.
FT   DISULFID   340    385       By similarity.
FT   DISULFID   384    393       By similarity.
FT   DISULFID   416    462       By similarity.
FT   DISULFID   461    472       By similarity.
FT   DISULFID   485    501       By similarity.
FT   DISULFID   500    511       By similarity.
FT   DISULFID   538    583       By similarity.
FT   DISULFID   582    591       By similarity.
SQ   SEQUENCE   608 AA;  68910 MW;  9EECAFDA86B1EF09 CRC64;

  Query Match             72.1%;  Score 2462;  DB 1;  Length 608;
  Best Local Similarity   74.4%;  Pred. No. 1.5e-147;
  Matches  435;  Conservative   78;  Mismatches   72;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              |:|||||:|||| |:|||:|  |||| |:||||:||:|:| ||| |||: || |||||||
Db         24 REAHKSEIAHRFNDVGEEHFIGLVLITFSQYLQKCPYEEHAKLVKEVTDLAKACVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||| :||||:| : :||:|||::|||| |:||||||||| |||| |:||   |||
Db         84 ANCDKSLHDIFGDKICALPSLRDTYGDVADCCEKKEPERNECFLHHKDDKPDLPPFARPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||:| ||||:|: |   ||||:|||||||||||::|::|||  ||||:|||| |||
Db        144 ADVLCKAFHDDEKAFFGHYLYEVARRHPYFYAPELLYYAQKYKAILTECCEAADKGACLT 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              |||| |  :   |:|::||:|||:||||:||:|||| |||||||||:| ::||:|||||
Db        204 PKLDALEGKSLISAAQERLRCASIQKFGDRAYKAWALVRLSQRFPKADFTDISKIVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||| |||||||||||||||||||:||:|::||| |||||:||:|||:|||  : ||| |
Db        264 KVHKECCHGDLLECADDRADLAKYMCEHQETISSHLKECCDKPILEKAHCIYGLHNDETP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              | ||::| :||| ||||||| ||||:||| |||||:|||||||||||||| | || ||:|
Db        324 AGLPAVAEEFVEDKDVCKNYEEAKDLFLGKFLYEYSRRHPDYSVVLLLRLGKAYEATLKK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              |||  ||| ||||| |||:|||:||:||:|||||:||||:| ||||||||||||||||
Db        384 CCATDDPHACYAKVLDEFQPLVDEPKNLVKQNCELYEQLGDYNFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||:||:|||||||||||||:|:|| |||||||||:||||||||||::||||||:|
Db        444 TPTLVEISRSLGKVGSKCCKHPEAERLPCVEDYLSVVLNRLCVLHEKTPVSEKVTKCCSE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              |||:|||||||  |||||||||||||||||||| | ||:|||||||||||||| |
Db        504 SLVDRRPCFSALGPDETYVPKEFNAETFTFHADICTLPETERKIKKQTALVELVKHKPHA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALG 644
              | :||| |: :| | ::||| |:||| ||| || ||| :|:| ||
Db        564 TNDQLKTVVGEFTALLDKCCSAEDKEACFAVEGPKLVESSKATLG 608


RESULT 15
Q5EG49_MICFO
ID   Q5EG49_MICFO PRELIMINARY;      PRT;   608 AA.
AC   Q5EG49;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Albumin.
OS   Microtus fortis fortis.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Arvicolinae; Microtus.
OX   NCBI_TaxID=311338;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Hu W.-X., Wu G.-J., Qin Z.-Q., Luo S.-Q.;
RT   "Albumin gene of Microtus fortis fortis liver.";
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY885264; AAW79112.1; -; mRNA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 2.
SQ   SEQUENCE   608 AA;  68309 MW;  060802BCD5E219D2 CRC64;

  Query Match             72.0%;  Score 2460;  DB 2;  Length 608;
  Best Local Similarity   75.9%;  Pred. No. 2e-147;
  Matches  443;  Conservative   66;  Mismatches   75;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              |||||||:|||: |||||: || |||| |||:||:||:|:|:|||||||:||| | |||||
Db         24 RDAHKSEIAHRYNDLGEKYFKGLVLITFAQHLQKCPYEEHIKLVNEVTDFAKACAADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              |||||||||||||||| :  | : | |||:||||||||||:|||:|||| |:||   ||||
Db         84 ENCDKSLHTLFGDKLCAIPNLGDNYAEMAECCAKQEPERNQCFLKHKDDKPSLPPFVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              :||||:| :|  :|:  ||: :|||||||||||::|::  |  ||||  |||||||
Db        144 AEVMCTSFQENPASFMGHYLHAVARRHPYFYAPELLYYAEKQSAIMTECCAEADKAACLG 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              |||| |:::   || |||||:||:|||||||||||||:||:||||:|||::|| ||||
Db        204 PKLDALKEKALLSSVKQRLKCSSLEKFGERAFKAWAVARMSQKFPKADFAEITKLATDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||  |||||||||| ||| :||||:|:|| :|||||   ||:||:|:|:||:|||::|||
Db        264 KVTQECCHGDLLECVDDRLELAKYMCDNQATISSKLHTCCDKPVLQKAHCLAEVDHDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              |||  || |||| |||||||||||||| ||||||||||||| ||||||| || ||||
Db        324 ADLTPLADDFVEDKDVCKNYAEAKDVFLGTFLYEYARRHPDYSVALLLRLAKKYEATLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||| |||| || ||||||:||||||:||:| ||||||:|||| ||||||||||:| ||||
Db        384 CCAEADPHACYGKVFDEFQPLVEEPKNLVKANCELFEKLGEYGFQNALLVRYTQKAPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              |||||||||||||||||:|||  ||| |:|| ||||| :|| ||||||||||::||||||
Db        444 TPTLVEVSRNLGKVGTKCCALPEADRLPCVEDYLSAILNHLCVLHEKTPVSEQVTKCCTG 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              |:| |||||||| ||| |||||| ||:|||||||| :||:|||||||| |||||||||
Db        504 SVVERRPCFSALPVDEHYVPKEFKAESFTFHADICTLPDKEKQIKKQTALAELVKHKPKA 563
```

```
Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL 643
              | :||| || :|:||:||||||||||| ||:||| |||| |||||
Db        564 TGDQLKTVMGEFSAFLEKCCKADDKEACFSEEGPKLVATSQAAL 607
```

Search completed: April 19, 2006, 12:08:51
Job time : 172.412 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:09:12 ; Search time 40.7706 Seconds
(without alignments)
1307.948 Million cell updates/sec

Title: US-10-775-180-447_COPY_30_674
Perfect score: 3417
Sequence: 1 HGEGTFTSDVSSYLEGQAAK.....TCFAEEGKKLVAASQAALGL 645

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3112.5	91.1	787	1	US-08-256-938-4	Sequence 4, Appli
2	3112.5	91.1	787	1	US-08-797-689-16	Sequence 16, Appl
3	3112.5	91.1	787	2	US-09-984-186-16	Sequence 16, Appl
4	3108	91.0	609	2	US-09-976-594-977	Sequence 977, App
5	3108	91.0	609	2	US-09-919-039-370	Sequence 370, App
6	3108	91.0	610	1	US-08-797-689-2	Sequence 2, Appli
7	3108	91.0	610	2	US-09-984-186-2	Sequence 2, Appli
8	3108	91.0	622	2	US-09-949-016-11170	Sequence 11170, A
9	3108	91.0	783	1	US-08-256-938-2	Sequence 2, Appli
10	3104	90.8	609	1	US-08-222-619-3	Sequence 3, Appli
11	3104	90.8	609	1	US-08-433-037-4	Sequence 4, Appli
12	3104	90.8	609	2	US-08-897-956A-2	Sequence 2, Appli
13	3104	90.8	609	4	PCT-US95-04075-3	Sequence 3, Appli
14	3103.5	90.8	978	2	US-08-897-956A-3	Sequence 3, Appli
15	3103	90.8	585	1	US-08-153-799-14	Sequence 14, Appl
16	3103	90.8	585	1	US-08-702-572-2	Sequence 2, Appli
17	3103	90.8	585	2	US-08-769-746-2	Sequence 2, Appli
18	3103	90.8	585	2	US-09-833-118A-18	Sequence 18, Appl
19	3103	90.8	585	2	US-09-832-929A-18	Sequence 18, Appl
20	3103	90.8	585	2	US-09-833-111A-18	Sequence 18, Appl
21	3093	90.5	585	1	US-08-448-196A-3	Sequence 3, Appli
22	3093	90.5	585	1	US-08-984-176-1	Sequence 1, Appli
23	2458.5	71.9	583	1	US-08-448-196A-5	Sequence 5, Appli
24	2450.5	71.7	583	1	US-08-448-196A-4	Sequence 4, Appli
25	2450.5	71.7	583	2	US-10-360-101-200	Sequence 200, App
26	2432.5	71.2	583	1	US-08-448-196A-6	Sequence 6, Appli
27	2426	71.0	584	1	US-08-448-196A-7	Sequence 7, Appli
28	2393.5	70.0	604	2	US-10-045-170A-1	Sequence 1, Appli
29	2389	69.9	582	1	US-08-134-638-1	Sequence 1, Appli
30	1256.5	36.8	609	1	US-08-222-619-4	Sequence 4, Appli
31	1256.5	36.8	609	2	US-09-976-594-456	Sequence 456, App
32	1256.5	36.8	609	4	PCT-US95-04075-4	Sequence 4, Appli
33	1256.5	36.8	612	2	US-09-949-016-11201	Sequence 11201, A
34	1213.5	35.5	609	2	US-09-186-949A-2	Sequence 2, Appli
35	1206.5	35.3	590	1	US-08-377-309-2	Sequence 2, Appli
36	1206.5	35.3	590	2	US-09-186-723-2	Sequence 2, Appli
37	1206.5	35.3	590	2	US-08-505-012-5	Sequence 5, Appli
38	1206.5	35.3	590	2	US-09-186-949A-3	Sequence 3, Appli
39	1206.5	35.3	590	2	US-08-758-757-2	Sequence 2, Appli
40	1206.5	35.3	590	2	US-09-187-978-2	Sequence 2, Appli
41	1206.5	35.3	590	2	US-10-115-701A-2	Sequence 2, Appli
42	1206.5	35.3	590	2	US-09-940-308A-2	Sequence 2, Appli
43	1206.5	35.3	590	2	US-09-940-308A-2	Sequence 2, Appli
44	1206.5	35.3	590	4	PCT-US96-00996-5	Sequence 5, Appli
45	1164.5	34.1	579	1	US-08-448-196A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne
; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: ST92007-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-256-938-4

Query Match 91.1%; Score 3112.5; DB 1; Length 787;
Best Local Similarity 94.6%; Pred. No. 3.9e-278;
Matches 596; Conservative 4; Mismatches 29; Indels 1; Gaps 1;

Qy 17 QAAKEFIAWLVKGRHGEGTFTSDVSSYLE-GQAAKEFIAWLVKGRDAHKSEVAHRFKDLG 75

```
Db        158 QGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQPGGGGDAHKSEVAHRFKDLG 217

Qy         76 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 135
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 277

Qy        136 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 195
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 337

Qy        196 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 397

Qy        256 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 315
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 457

Qy        316 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 375
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 517

Qy        376 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 435
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 577

Qy        436 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 495
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 637

Qy        496 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 555
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 697

Qy        556 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 615
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 757

Qy        616 EKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||
Db        758 EKCCKADDKETCFAEEGKKLVAASQAALGL 787


RESULT 2
US-08-797-689-16
; Sequence 16, Application US/08797689
; Patent No. 5876969
;  GENERAL INFORMATION:
;    APPLICANT:  Fleer, Reinhard
;    APPLICANT:  Fournier, Alain
;    APPLICANT:  Guitton, Jean-Dominique
;    APPLICANT:  Jung, Gerard
;    APPLICANT:  Yeh, Patrice
;    TITLE OF INVENTION:  NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;    TITLE OF INVENTION:  PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;    TITLE OF INVENTION:  CONTAINING SAID POLYPEPTIDES
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.1
;      SOFTWARE:  Word 5.1 (PatentIn)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/797,689
;      FILING DATE:  31-JAN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/256,927
;      FILING DATE:  28-JUL-1994
;      APPLICATION NUMBER:  FR 92/01064
;      FILING DATE:  31-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00085
;      FILING DATE:  28-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith Ph.D., Julie K.
;      REGISTRATION NUMBER:  P-38,619
;      REFERENCE/DOCKET NUMBER:  ST92006-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3839
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  787 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-797-689-16

  Query Match             91.1%;  Score 3112.5;  DB 1;  Length 787;
  Best Local Similarity   94.6%;  Pred. No. 3.9e-278;
  Matches  596;  Conservative    4;  Mismatches   29;  Indels    1;  Gaps    1;

Qy         17 QAAKEFIAWLVKGRHGEGTFTSDVSSYLE-GQAAKEFIAWLVKGRDAHKSEVAHRFKDLG 75
              | |    |   : | |     | : |:||      :|    | ||||||||||||||| 
Db        158 QGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQPGGGGDAHKSEVAHRFKDLG 217

Qy         76 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 135
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 277

Qy        136 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 195
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 337

Qy        196 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 397

Qy        256 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 315
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 457

Qy        316 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 375
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 517

Qy        376 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 435
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 577

Qy        436 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 495
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 637

Qy        496 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 555
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 697

Qy        556 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 615
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 757

Qy        616 EKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||
Db        758 EKCCKADDKETCFAEEGKKLVAASQAALGL 787


RESULT 3
```

```
US-09-984-186-16
; Sequence 16, Application US/09984186
; Patent No. 6686179
;   GENERAL INFORMATION:
;        APPLICANT: Fleer, Reinhard
;                   Fournier, Alain
;                   Guitton, Jean-Dominique
;                   Jung, Gerard
;                   Yeh, Patrice
;        TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                            PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                            CONTAINING SAID POLYPEPTIDES
;        NUMBER OF SEQUENCES: 36
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rhone-Poulenc Rorer Inc.
;             STREET: 500 Arcola Road, 3C43
;             CITY: Collegeville
;             STATE: PA
;             COUNTRY: USA
;             ZIP: 19426
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: Macintosh
;             OPERATING SYSTEM: System 7.1
;             SOFTWARE: Word 5.1 (PatentIn)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/984,186
;             FILING DATE: 29-Oct-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/797,689
;             FILING DATE: 31-JAN-1997
;             APPLICATION NUMBER: US 08/256,927
;             FILING DATE: 28-JUL-1994
;             APPLICATION NUMBER: FR 92/01064
;             FILING DATE: 31-JAN-1992
;             APPLICATION NUMBER: PCT/FR93/00085
;             FILING DATE: 28-JAN-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Smith Ph.D., Julie K.
;             REGISTRATION NUMBER: P-38,619
;             REFERENCE/DOCKET NUMBER: ST92006-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (610) 454-3839
;             TELEFAX: (610) 454-3808
;   INFORMATION FOR SEQ ID NO: 16:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 787 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-984-186-16

  Query Match             91.1%;  Score 3112.5;  DB 2;  Length 787;
  Best Local Similarity   94.6%;  Pred. No. 3.9e-278;
  Matches  596;  Conservative    4;  Mismatches   29;  Indels    1;  Gaps    1;

Qy         17 QAAKEFIAWLVKGRHGEGTFTSDVSSYLE-GQAAKEFIAWLVKGRDAHKSEVAHRFKDLG 75
              | |    |   : | |      | : |:||        :|    | ||||||||||||||
Db        158 QGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQPGGGGDAHKSEVAHRFKDLG 217

Qy         76 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 EENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 277

Qy        136 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFL 337

Qy        196 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 KKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAK 397

Qy        256 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 QRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECAD 457

Qy        316 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 DRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 517

Qy        376 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 CKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD 577

Qy        436 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGS 637

Qy        496 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 KCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDE 697

Qy        556 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFV 757

Qy        616 EKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||
Db        758 EKCCKADDKETCFAEEGKKLVAASQAALGL 787


RESULT 4
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 977
;   LENGTH: 609
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

  Query Match             91.0%;  Score 3108;  DB 2;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-278;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263
```

```
Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

RESULT 5
US-09-919-039-370
; Sequence 370, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6727066 088957CD1
US-09-919-039-370

Query Match 91.0%; Score 3108; DB 2; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.9e-278;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

RESULT 6
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-2

```
  Query Match             91.0%;  Score 3108;  DB 1;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-278;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy         120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy         180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy         240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy         300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy         360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy         420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy         480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy         540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy         600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
               |||||||||||||||||||||||||||||||||||||||||||||
Db         564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 7
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
;   GENERAL INFORMATION:
;        APPLICANT: Fleer, Reinhard
;                   Fournier, Alain
;                   Guitton, Jean-Dominique
;                   Jung, Gerard
;                   Yeh, Patrice
;        TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                            PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                            CONTAINING SAID POLYPEPTIDES
;        NUMBER OF SEQUENCES: 36
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rhone-Poulenc Rorer Inc.
;             STREET: 500 Arcola Road, 3C43
;             CITY: Collegeville
;             STATE: PA
;             COUNTRY: USA
;             ZIP: 19426
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: Macintosh
;             OPERATING SYSTEM: System 7.1
;             SOFTWARE: Word 5.1 (PatentIn)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/984,186
;             FILING DATE: 29-Oct-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/797,689
;             FILING DATE: 31-JAN-1997
;             APPLICATION NUMBER: US 08/256,927
;             FILING DATE: 28-JUL-1994
;             APPLICATION NUMBER: FR 92/01064
;             FILING DATE: 31-JAN-1992
;             APPLICATION NUMBER: PCT/FR93/00085
;             FILING DATE: 28-JAN-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Smith Ph.D., Julie K.
;             REGISTRATION NUMBER: P-38,619
;             REFERENCE/DOCKET NUMBER: ST92006-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (610) 454-3839
;             TELEFAX: (610) 454-3808
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 610 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

  Query Match             91.0%;  Score 3108;  DB 2;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-278;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy         120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy         180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy         240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy         300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy         360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy         420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy         480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy         540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy         600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
               |||||||||||||||||||||||||||||||||||||||||||||
Db         564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

```
RESULT 8
US-09-949-016-11170
; Sequence 11170, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11170
;   LENGTH: 622
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-11170

  Query Match             91.0%;  Score 3108;  DB 2;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-278;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 96

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 156

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 216

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 276

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 336

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 396

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 456

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 516

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 576

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        577 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 622


RESULT 9
US-08-256-938-2
; Sequence 2, Application US/08256938
; Patent No. 5665863
;  GENERAL INFORMATION:
;    APPLICANT:  Yeh, Patrice
;    TITLE OF INVENTION:  NEW POLYPEPTIDES HAVING GRANULOCYTE
;    TITLE OF INVENTION:  COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
;    TITLE OF INVENTION:  PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.1
;      SOFTWARE:  Word 5.0 (PatentIn)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/256,938
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 92/01065
;      FILING DATE:  31-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goodman, Rosanne
;      REGISTRATION NUMBER:  32,534
;      REFERENCE/DOCKET NUMBER:  ST92007-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3817
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  783 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-256-938-2

  Query Match             91.0%;  Score 3108;  DB 1;  Length 783;
  Best Local Similarity   100.0%;  Pred. No. 1e-277;
  Matches  586;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443
```

```
Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 10
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
;  GENERAL INFORMATION:
;    APPLICANT:  Lichenstein, Henri
;    APPLICANT:  Lyons, David
;    APPLICANT:  Wurfel, Mark
;    APPLICANT:  Wright, Samuel
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/222,619
;      FILING DATE:
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  609 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-222-619-3

  Query Match             90.8%;  Score 3104;  DB 1;  Length 609;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-277;
  Matches  585;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFKQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 11
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
;  GENERAL INFORMATION:
;    APPLICANT:  Sreekrishna, Kotikanyadan
;    APPLICANT:  Barr, Kathryn A.
;    APPLICANT:  Brierley, Russell A.
;    APPLICANT:  Thill, Gregory P.
;    APPLICANT:  Tschopp, Juerg F.
;    TITLE OF INVENTION:  EXPRESSION OF HUMAN SERUM ALBUMIN IN
;    TITLE OF INVENTION:  PICHIA PASTORIS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/433,037
;      FILING DATE:  03-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9108Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  609 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-433-037-4

  Query Match             90.8%;  Score 3104;  DB 1;  Length 609;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-277;
  Matches  585;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKVSSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 12
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
;  APPLICANT:  Mary Ellen Digan
;  APPLICANT:  Philip Lake
;  APPLICANT:  Hermann Gram
;  TITLE OF INVENTION:  Fusion Polypeptides
;  FILE REFERENCE:  600-7244/CPA
;  CURRENT APPLICATION NUMBER: US/08/897,956A
;  CURRENT FILING DATE:  1997-07-21
;  PRIOR APPLICATION NUMBER:  60/022,689
;  PRIOR FILING DATE:  1996-07-26
;  NUMBER OF SEQ ID NOS:  38
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 609
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-08-897-956A-2

  Query Match             90.8%;  Score 3104;  DB 2;  Length 609;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-277;
  Matches  585;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83

Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFKQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609


RESULT 13
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
;  GENERAL INFORMATION:
;    APPLICANT:  AMGEN INC.
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04075
;      FILING DATE:
;      CLASSIFICATION:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  609 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
PCT-US95-04075-3

  Query Match             90.8%;  Score 3104;  DB 4;  Length 609;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-277;
  Matches  585;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 RDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESA 83
```

```
Qy        120 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPE 143

Qy        180 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLL 203

Qy        240 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 PKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 263

Qy        300 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 323

Qy        360 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEK 383

Qy        420 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVS 479
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        384 CCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFKQLGEYKFQNALLVRYTKKVPQVS 443

Qy        480 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTE 503

Qy        540 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 SLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKA 563

Qy        600 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        564 TKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 609
```

RESULT 14
US-08-897-956A-3
; Sequence 3, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 978
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion polypeptide
US-08-897-956A-3

Query Match 90.8%; Score 3103.5; DB 2; Length 978;
Best Local Similarity 98.8%; Pred. No. 3.6e-277;
Matches 586; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

```
Qy         55 WLVK---GRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAK 111
              ||    | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 WLASGGGGSDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAK 262

Qy        112 TCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPN 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 TCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPN 322

Qy        172 LPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQA 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 LPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQA 382

Qy        232 ADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEV 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 ADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEV 442

Qy        292 SKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 SKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA 502

Qy        352 EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAK 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAK 562

Qy        412 TYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRY 471
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        563 TYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFKQLGEYKFQNALLVRY 622

Qy        472 TKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSD 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 TKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSD 682

Qy        532 RVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVE 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 RVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVE 742

Qy        592 LVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALG 644
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 LVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALG 795
```

RESULT 15
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain

```
;      REGISTRATION NUMBER:  24864
;      REFERENCE/DOCKET NUMBER:  92H832
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 665 2400
;      TELEFAX:  (908) 771 6159
;      TELEX:  219484
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  369..419
;      OTHER INFORMATION:  /note= "Alternative C-termini of
;      OTHER INFORMATION:  HSA(1-n)"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..585
;      OTHER INFORMATION:  /note= "Amino acid sequence of
;      OTHER INFORMATION:  natural HSA"
US-08-153-799-14

  Query Match             90.8%;  Score 3103;  DB 1;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-277;
  Matches  585;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 60

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 120

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 180

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 360

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 420

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        541 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 585

Search completed: April 19, 2006, 12:11:36
```

Job time : 42.7706 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:29:13 ; Search time 136.055 Seconds
(without alignments)
1980.821 Million cell updates/sec

Title: US-10-775-180-447_COPY_30_674
Perfect score: 3417
Sequence: 1 HGEGTFTSDVSSYLEGQAAK.....TCFAEEGKKLVAASQAALGL 645

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3417	100.0	662	5	US-10-775-180-611	Sequence 611, App
2	3417	100.0	662	5	US-10-775-204-1623	Sequence 1623, Ap
3	3417	100.0	663	5	US-10-775-180-600	Sequence 600, App
4	3417	100.0	663	5	US-10-775-204-1609	Sequence 1609, Ap
5	3417	100.0	664	5	US-10-775-180-598	Sequence 598, App
6	3417	100.0	664	5	US-10-775-204-1607	Sequence 1607, Ap
7	3417	100.0	668	5	US-10-775-180-609	Sequence 609, App
8	3417	100.0	668	5	US-10-775-204-1621	Sequence 1621, Ap
9	3417	100.0	669	5	US-10-775-180-419	Sequence 419, App
10	3417	100.0	669	5	US-10-775-204-1231	Sequence 1231, Ap
11	3417	100.0	674	5	US-10-775-180-447	Sequence 447, App
12	3417	100.0	674	5	US-10-775-204-1280	Sequence 1280, Ap
13	3417	100.0	730	5	US-10-775-180-610	Sequence 610, App
14	3417	100.0	730	5	US-10-775-204-1622	Sequence 1622, Ap
15	3411	99.8	662	5	US-10-775-180-614	Sequence 614, App
16	3411	99.8	662	5	US-10-775-204-1626	Sequence 1626, Ap
17	3411	99.8	663	5	US-10-775-180-601	Sequence 601, App
18	3411	99.8	663	5	US-10-775-204-1610	Sequence 1610, Ap
19	3411	99.8	664	5	US-10-775-180-599	Sequence 599, App
20	3411	99.8	664	5	US-10-775-204-1608	Sequence 1608, Ap
21	3411	99.8	668	5	US-10-775-180-613	Sequence 613, App
22	3411	99.8	668	5	US-10-775-204-1625	Sequence 1625, Ap
23	3411	99.8	669	5	US-10-775-180-425	Sequence 425, App
24	3411	99.8	669	5	US-10-775-204-1237	Sequence 1237, Ap
25	3411	99.8	730	5	US-10-775-180-612	Sequence 612, App
26	3411	99.8	730	5	US-10-775-204-1624	Sequence 1624, Ap
27	3405	99.6	669	5	US-10-775-180-420	Sequence 420, App
28	3405	99.6	669	5	US-10-775-180-421	Sequence 421, App
29	3405	99.6	669	5	US-10-775-180-423	Sequence 423, App
30	3405	99.6	669	5	US-10-775-180-424	Sequence 424, App
31	3405	99.6	669	5	US-10-775-204-1232	Sequence 1232, Ap
32	3405	99.6	669	5	US-10-775-204-1233	Sequence 1233, Ap
33	3405	99.6	669	5	US-10-775-204-1235	Sequence 1235, Ap
34	3405	99.6	669	5	US-10-775-204-1236	Sequence 1236, Ap
35	3397	99.4	667	5	US-10-775-180-422	Sequence 422, App
36	3397	99.4	667	5	US-10-775-204-1234	Sequence 1234, Ap
37	3265	95.6	639	5	US-10-775-180-131	Sequence 131, App
38	3265	95.6	639	5	US-10-775-204-417	Sequence 417, App
39	3259	95.4	639	5	US-10-775-180-129	Sequence 129, App
40	3259	95.4	639	5	US-10-775-204-414	Sequence 414, App
41	3259	95.4	700	5	US-10-775-204-1620	Sequence 1620, Ap
42	3250.5	95.1	654	5	US-10-775-180-574	Sequence 574, App
43	3250.5	95.1	654	5	US-10-775-204-1559	Sequence 1559, Ap
44	3248	95.1	655	5	US-10-775-180-623	Sequence 623, App
45	3248	95.1	655	5	US-10-775-204-1640	Sequence 1640, Ap

ALIGNMENTS

RESULT 1
US-10-775-180-611
; Sequence 611, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 611
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-180-611

Query Match 100.0%; Score 3417; DB 5; Length 662;
Best Local Similarity 100.0%; Pred. No. 3.6e-259;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 77

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         78 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 137

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 197

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 257

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 317

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 377

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 437

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 497

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 557

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 617

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        618 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 662


RESULT 2
US-10-775-204-1623
; Sequence 1623, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1623
;   LENGTH: 662
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1623

  Query Match           100.0%;  Score 3417;  DB 5;  Length 662;
  Best Local Similarity  100.0%;  Pred. No. 3.6e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 77

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 137

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 197

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 257

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 317

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 377

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 437

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 497

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 557

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 617

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        618 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 662


RESULT 3
US-10-775-180-600
; Sequence 600, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
```

```
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 600
;   LENGTH: 663
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-600

  Query Match           100.0%;  Score 3417;  DB 5;  Length 663;
  Best Local Similarity  100.0%;  Pred. No. 3.6e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 78

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 138

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 198

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 258

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 318

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 378

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 438

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 498

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 558

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 618

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        619 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 663
```

```
RESULT 4
US-10-775-204-1609
; Sequence 1609, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1609
;   LENGTH: 663
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1609

  Query Match           100.0%;  Score 3417;  DB 5;  Length 663;
  Best Local Similarity  100.0%;  Pred. No. 3.6e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 78

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 138

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 198

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 258

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 318

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 378

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 438
```

```
Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 498

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 558

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 618

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        619 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 663
```

RESULT 5
US-10-775-180-598
; Sequence 598, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 598
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-180-598

Query Match 100.0%; Score 3417; DB 5; Length 664;
Best Local Similarity 100.0%; Pred. No. 3.6e-259;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 79

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 139

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 199

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 259

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 319

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 379

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 439

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 499

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 559

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 619

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        620 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 664
```

RESULT 6
US-10-775-204-1607
; Sequence 1607, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1607
; LENGTH: 664
; TYPE: PRT

```
;   ORGANISM: Homo sapiens
US-10-775-204-1607

  Query Match             100.0%;  Score 3417;  DB 5;  Length 664;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 79

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 139

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 199

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 259

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 319

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 379

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 439

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 499

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 559

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 619

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        620 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 664


RESULT 7
US-10-775-180-609
; Sequence 609, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 609
;   LENGTH: 668
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-609

  Query Match             100.0%;  Score 3417;  DB 5;  Length 668;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 83

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 143

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 203

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 263

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 323

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 383

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 443

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 503

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 563

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 623

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        624 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 668


RESULT 8
US-10-775-204-1621
; Sequence 1621, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1621
;   LENGTH: 668
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1621

  Query Match             100.0%;  Score 3417;  DB 5;  Length 668;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 83

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 143

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 203

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 263

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 323

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 383

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 443

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 503

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 563

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 623

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        624 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 668


RESULT 9
US-10-775-180-419
; Sequence 419, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 419
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-419

  Query Match             100.0%;  Score 3417;  DB 5;  Length 669;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 84

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 144

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 204

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 264
```

```
Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 324

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 384

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 444

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 504

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 564

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 624

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        625 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 669
```

RESULT 10
US-10-775-204-1231
; Sequence 1231, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1231
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-1231

Query Match 100.0%; Score 3417; DB 5; Length 669;
Best Local Similarity 100.0%; Pred. No. 3.6e-259;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 84

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 144

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 204

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 264

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 324

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 384

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 444

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 504

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 564

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 624

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        625 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 669
```

RESULT 11
US-10-775-180-447
; Sequence 447, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611

```
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 447
;   LENGTH: 674
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-447

  Query Match             100.0%;  Score 3417;  DB 5;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 149

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 209

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 269

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 329

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 389

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 449

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 509

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 569

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 629

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        630 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674


RESULT 12
US-10-775-204-1280
; Sequence 1280, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1280
;   LENGTH: 674
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1280

  Query Match             100.0%;  Score 3417;  DB 5;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 149

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 209

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 269

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 329

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 389

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 449

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 509

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 569

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 629

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        630 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674


RESULT 13
US-10-775-180-610
; Sequence 610, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 610
;   LENGTH: 730
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-610

  Query Match             100.0%;  Score 3417;  DB 5;  Length 730;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 145

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 205

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 265

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 325

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 385

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 445

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 505

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 565

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 625

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 685

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        686 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 730


RESULT 14
US-10-775-204-1622
; Sequence 1622, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1622
;   LENGTH: 730
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1622

  Query Match             100.0%;  Score 3417;  DB 5;  Length 730;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-259;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 145
```

```
Qy          61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         146 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 205

Qy         121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         206 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 265

Qy         181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         266 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 325

Qy         241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         326 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 385

Qy         301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         386 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 445

Qy         361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         446 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 505

Qy         421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         506 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 565

Qy         481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         566 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 625

Qy         541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         626 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 685

Qy         601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
               |||||||||||||||||||||||||||||||||||||||||||||
Db         686 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 730


RESULT 15
US-10-775-180-614
; Sequence 614, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 614
;   LENGTH: 662
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-614

  Query Match             99.8%;  Score 3411;  DB 5;  Length 662;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-258;
  Matches  644;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
               |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          18 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 77

Qy          61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          78 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 137

Qy         121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         138 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 197

Qy         181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         198 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 257

Qy         241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 317

Qy         301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         318 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 377

Qy         361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         378 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 437

Qy         421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         438 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 497

Qy         481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         498 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 557

Qy         541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         558 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 617

Qy         601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
               |||||||||||||||||||||||||||||||||||||||||||||
Db         618 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 662


Search completed: April 19, 2006, 12:35:50
Job time : 138.055 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:31:03 ; Search time 21.5305 Seconds
(without alignments)
1318.215 Million cell updates/sec

Title: US-10-775-180-447_COPY_30_674
Perfect score: 3417
Sequence: 1.HGEGTFTSDVSSYLEGQAAK.....TCFAEEGKKLVAASQAALGL 645

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225428 seqs, 44002918 residues

Total number of hits satisfying chosen parameters: 225428

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /SIDS5/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /SIDS5/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /SIDS5/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /SIDS5/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /SIDS5/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /SIDS5/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /SIDS5/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /SIDS5/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3417	100.0	674	7	US-11-175-690-206	Sequence 206, App
2	3417	100.0	915	7	US-11-175-690-208	Sequence 208, App
3	3254	95.2	646	7	US-11-175-690-223	Sequence 223, App
4	3253.5	95.2	647	7	US-11-175-690-212	Sequence 212, App
5	3253	95.2	648	7	US-11-175-690-214	Sequence 214, App
6	3252.5	95.2	649	7	US-11-175-690-213	Sequence 213, App
7	3252	95.2	650	7	US-11-175-690-209	Sequence 209, App
8	3251.5	95.2	651	7	US-11-175-690-224	Sequence 224, App
9	3251	95.1	652	7	US-11-175-690-218	Sequence 218, App
10	3250.5	95.1	653	7	US-11-175-690-215	Sequence 215, App
11	3250	95.1	654	7	US-11-175-690-219	Sequence 219, App
12	3250	95.1	658	7	US-11-175-690-210	Sequence 210, App
13	3249.5	95.1	655	7	US-11-175-690-220	Sequence 220, App
14	3249	95.1	656	7	US-11-175-690-225	Sequence 225, App
15	3248.5	95.1	657	7	US-11-175-690-216	Sequence 216, App
16	3248.5	95.1	657	7	US-11-175-690-303	Sequence 303, App
17	3247.5	95.0	659	7	US-11-175-690-221	Sequence 221, App
18	3183.5	93.2	646	7	US-11-175-690-276	Sequence 276, App
19	3183.5	93.2	678	7	US-11-175-690-274	Sequence 274, App
20	3173.5	92.9	642	7	US-11-175-690-238	Sequence 238, App
21	3167.5	92.7	642	7	US-11-175-690-233	Sequence 233, App
22	3154	92.3	647	7	US-11-175-690-242	Sequence 242, App
23	3151.5	92.2	636	7	US-11-175-690-268	Sequence 268, App
24	3150.5	92.2	636	7	US-11-175-690-278	Sequence 278, App
25	3146.5	92.1	636	7	US-11-175-690-240	Sequence 240, App
26	3135	91.7	688	7	US-11-175-690-198	Sequence 198, App
27	3135	91.7	693	7	US-11-175-690-199	Sequence 199, App
28	3131.5	91.6	629	7	US-11-175-690-562	Sequence 562, App
29	3127	91.5	637	7	US-11-175-690-266	Sequence 266, App
30	3125	91.5	672	7	US-11-175-690-200	Sequence 200, App
31	3118	91.2	728	7	US-11-175-690-244	Sequence 244, App
32	3118	91.2	728	7	US-11-175-690-246	Sequence 246, App
33	3118	91.2	728	7	US-11-175-690-248	Sequence 248, App
34	3112.5	91.1	634	7	US-11-175-690-207	Sequence 207, App
35	3109.5	91.0	638	7	US-11-175-690-229	Sequence 229, App
36	3108.5	91.0	742	7	US-11-175-690-525	Sequence 525, App
37	3108	91.0	609	7	US-11-175-690-3	Sequence 3, Appli
38	3108	91.0	629	7	US-11-175-690-561	Sequence 561, App
39	3108	91.0	634	7	US-11-175-690-279	Sequence 279, App
40	3108	91.0	636	7	US-11-175-690-239	Sequence 239, App
41	3108	91.0	636	7	US-11-175-690-267	Sequence 267, App
42	3108	91.0	636	7	US-11-175-690-277	Sequence 277, App
43	3108	91.0	637	7	US-11-175-690-265	Sequence 265, App
44	3108	91.0	637	7	US-11-175-690-557	Sequence 557, App
45	3108	91.0	638	7	US-11-175-690-559	Sequence 559, App

ALIGNMENTS

RESULT 1
US-11-175-690-206
; Sequence 206, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 206
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-206

Query Match 100.0%; Score 3417; DB 7; Length 674;
Best Local Similarity 100.0%; Pred. No. 6.8e-262;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 149

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        150 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 209

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 269

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 329

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 389

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 449

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 509

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 569

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 629

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        630 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674


RESULT 2
US-11-175-690-208
; Sequence 208, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 208
;   LENGTH: 915
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-208

  Query Match           100.0%;  Score 3417;  DB 7;  Length 915;
  Best Local Similarity  100.0%;  Pred. No. 1e-261;
  Matches  645;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 149

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 209

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 269

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 329

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 389

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 449

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 509

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 569

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 629

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        630 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 674


RESULT 3
US-11-175-690-223
; Sequence 223, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
```

```
; SEQ ID NO 223
;   LENGTH: 646
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-223

  Query Match             95.2%;  Score 3254;  DB 7;  Length 646;
  Best Local Similarity   99.7%;  Pred. No. 4.8e-249;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR--DAHKSEVAHRFKDLGEENFKALVLIAF 87
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDADAHKSEVAHRFKDLGEENFKALVLIAF 88

Qy         88 AQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEM 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 AQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEM 148

Qy        148 ADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHP 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 ADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHP 208

Qy        208 YFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFG 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 YFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFG 268

Qy        268 ERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICEN 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 ERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICEN 328

Qy        328 QDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFL 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 QDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFL 388

Qy        388 GMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNL 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNL 448

Qy        448 IKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMP 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 IKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMP 508

Qy        508 CAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETF 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 CAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETF 568

Qy        568 TFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETC 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 TFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETC 628

Qy        628 FAEEGKKLVAASQAALGL 645
              ||||||||||||||||||
Db        629 FAEEGKKLVAASQAALGL 646


RESULT 4
US-11-175-690-212
; Sequence 212, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 212
;   LENGTH: 647
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-212

  Query Match             95.2%;  Score 3253.5;  DB 7;  Length 647;
  Best Local Similarity   99.5%;  Pred. No. 5.3e-249;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR---DAHKSEVAHRFKDLGEENFKALVLIA 86
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHDAHKSEVAHRFKDLGEENFKALVLIA 88

Qy         87 FAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGE 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 FAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGE 148

Qy        147 MADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRH 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 MADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRH 208

Qy        207 PYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 PYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF 268

Qy        267 GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICE 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICE 328

Qy        327 NQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVF 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 NQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVF 388

Qy        387 LGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQN 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 LGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQN 448

Qy        447 LIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRM 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 LIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRM 508

Qy        507 PCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAET 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 PCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAET 568

Qy        567 FTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKET 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 FTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKET 628

Qy        627 CFAEEGKKLVAASQAALGL 645
              ||||||||||||||||||
Db        629 CFAEEGKKLVAASQAALGL 647


RESULT 5
US-11-175-690-214
; Sequence 214, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
```

```
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 214
;   LENGTH: 648
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-214

  Query Match             95.2%;  Score 3253;  DB 7;  Length 648;
  Best Local Similarity   99.4%;  Pred. No. 5.8e-249;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR----DAHKSEVAHRFKDLGEENFKALVLI 85
              ||||||||||||||||||||||||||||||    |||||||||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKDAHKSEVAHRFKDLGEENFKALVLI 88

Qy         86 AFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYG 145
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 AFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYG 148

Qy        146 EMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARR 205
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 EMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARR 208

Qy        206 HPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQK 265
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 HPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQK 268

Qy        266 FGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYIC 325
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 FGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYIC 328

Qy        326 ENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDV 385
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDV 388

Qy        386 FLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQ 445
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 FLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQ 448

Qy        446 NLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKR 505
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 NLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKR 508

Qy        506 MPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAE 565
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 MPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAE 568

Qy        566 TFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKE 625
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 TFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKE 628

Qy        626 TCFAEEGKKLVAASQAALGL 645
              ||||||||||||||||||||
Db        629 TCFAEEGKKLVAASQAALGL 648
```

```
RESULT 6
US-11-175-690-213
; Sequence 213, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 213
;   LENGTH: 649
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-213

  Query Match             95.2%;  Score 3252.5;  DB 7;  Length 649;
  Best Local Similarity   99.2%;  Pred. No. 6.4e-249;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR-----DAHKSEVAHRFKDLGEENFKALVL 84
              ||||||||||||||||||||||||||||||     |||||||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSDAHKSEVAHRFKDLGEENFKALVL 88

Qy         85 IAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETY 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 IAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETY 148

Qy        145 GEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIAR 204
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 GEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIAR 208

Qy        205 RHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQ 264
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 RHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQ 268

Qy        265 KFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYI 324
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 KFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYI 328

Qy        325 CENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKD 384
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKD 388

Qy        385 VFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEP 444
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 VFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEP 448

Qy        445 QNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK 504
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 QNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK 508

Qy        505 RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNA 564
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNA 568

Qy        565 ETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDK 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 ETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDK 628

Qy        625 ETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||
Db        629 ETCFAEEGKKLVAASQAALGL 649


RESULT 7
US-11-175-690-209
; Sequence 209, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 209
;   LENGTH: 650
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-209

  Query Match             95.2%;  Score 3252;  DB 7;  Length 650;
  Best Local Similarity   99.0%;  Pred. No. 7e-249;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    6;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR------DAHKSEVAHRFKDLGEENFKALV 83
              |||||||||||||||||||||||||||||||      |||||||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEDAHKSEVAHRFKDLGEENFKALV 88

Qy         84 LIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRET 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 LIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRET 148

Qy        144 YGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 YGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA 208

Qy        204 RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASL 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASL 268

Qy        264 QKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKY 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 QKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKY 328

Qy        324 ICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAK 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAK 388

Qy        384 DVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEE 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 DVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEE 448

Qy        444 PQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEA 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 PQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEA 508

Qy        504 KRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFN 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 KRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFN 568

Qy        564 AETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADD 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 AETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADD 628

Qy        624 KETCFAEEGKKLVAASQAALGL 645
              ||||||||||||||||||||||
Db        629 KETCFAEEGKKLVAASQAALGL 650


RESULT 8
US-11-175-690-224
; Sequence 224, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 224
;   LENGTH: 651
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-224

  Query Match             95.2%;  Score 3251.5;  DB 7;  Length 651;
  Best Local Similarity   98.9%;  Pred. No. 7.7e-249;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    7;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR-------DAHKSEVAHRFKDLGEENFKAL 82
              |||||||||||||||||||||||||||||||       |||||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVDAHKSEVAHRFKDLGEENFKAL 88

Qy         83 VLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRE 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 VLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRE 148

Qy        143 TYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEI 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 TYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEI 208

Qy        203 ARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCAS 262
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 ARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCAS 268

Qy        263 LQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAK 322
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 LQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAK 328

Qy        323 YICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEA 382
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 YICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEA 388

Qy        383 KDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVE 442
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 KDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVE 448

Qy        443 EPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPE 502
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 EPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPE 508

Qy        503 AKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEF 562
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEF 568

Qy        563 NAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKAD 622
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 NAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKAD 628

Qy        623 DKETCFAEEGKKLVAASQAALGL 645
                 |||||||||||||||||||||||
Db        629 DKETCFAEEGKKLVAASQAALGL 651


RESULT 9
US-11-175-690-218
; Sequence 218, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 218
;   LENGTH: 652
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-218

  Query Match             95.1%;  Score 3251;  DB 7;  Length 652;
  Best Local Similarity   98.7%;  Pred. No. 8.4e-249;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR--------DAHKSEVAHRFKDLGEENFKA 81
                 ||||||||||||||||||||||||||||||        ||||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVADAHKSEVAHRFKDLGEENFKA 88

Qy         82 LVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLR 141
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 LVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLR 148

Qy        142 ETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYE 201
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 ETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYE 208

Qy        202 IARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCA 261
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 IARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCA 268

Qy        262 SLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLA 321
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 SLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLA 328

Qy        322 KYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAE 381
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 KYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAE 388

Qy        382 AKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLV 441
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 AKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLV 448

Qy        442 EEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHP 501
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 EEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHP 508

Qy        502 EAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKE 561
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 EAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKE 568

Qy        562 FNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKA 621
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 FNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKA 628

Qy        622 DDKETCFAEEGKKLVAASQAALGL 645
                 |||||||||||||||||||||||
Db        629 DDKETCFAEEGKKLVAASQAALGL 652


RESULT 10
US-11-175-690-215
; Sequence 215, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 215
;   LENGTH: 653
;   TYPE: PRT
```

```
;   ORGANISM: Homo sapiens
US-11-175-690-215

  Query Match             95.1%;  Score 3250.5;  DB 7;  Length 653;
  Best Local Similarity   98.6%;  Pred. No. 9.2e-249;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    9;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR---------DAHKSEVAHRFKDLGEENFK 80
              ||||||||||||||||||||||||||||||         |||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHDAHKSEVAHRFKDLGEENFK 88

Qy         81 ALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATL 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 ALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATL 148

Qy        141 RETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLY 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 RETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLY 208

Qy        201 EIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKC 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 EIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKC 268

Qy        261 ASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADL 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 ASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADL 328

Qy        321 AKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYA 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYA 388

Qy        381 EAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPL 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 EAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPL 448

Qy        441 VEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 VEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH 508

Qy        501 PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPK 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPK 568

Qy        561 EFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCK 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 EFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCK 628

Qy        621 ADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||
Db        629 ADDKETCFAEEGKKLVAASQAALGL 653


RESULT 11
US-11-175-690-219
; Sequence 219, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 219
;   LENGTH: 654
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-219

  Query Match             95.1%;  Score 3250;  DB 7;  Length 654;
  Best Local Similarity   98.4%;  Pred. No. 1e-248;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels   10;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKG----------RDAHKSEVAHRFKDLGEENF 79
              ||||||||||||||||||||||||||||||          ||||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHDAHKSEVAHRFKDLGEENF 88

Qy         80 KALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVAT 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 KALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVAT 148

Qy        140 LRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 LRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL 208

Qy        200 YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLK 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLK 268

Qy        260 CASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRAD 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 CASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRAD 328

Qy        320 LAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 LAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY 388

Qy        380 AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKP 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKP 448

Qy        440 LVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCK 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 LVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCK 508

Qy        500 HPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVP 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 HPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVP 568

Qy        560 KEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCC 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 KEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCC 628

Qy        620 KADDKETCFAEEGKKLVAASQAALGL 645
              ||||||||||||||||||||||||||
Db        629 KADDKETCFAEEGKKLVAASQAALGL 654


RESULT 12
US-11-175-690-210
; Sequence 210, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
```

```
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 210
;   LENGTH: 658
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-210

  Query Match             95.1%;  Score 3250;  DB 7;  Length 658;
  Best Local Similarity   96.0%;  Pred. No. 1e-248;
  Matches  619;  Conservative    1;  Mismatches    9;  Indels   16;  Gaps    1;

Qy          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR----------------DAHKSEVAHRFKDL 73

Qy         61 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAE 133

Qy        121 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEV 193

Qy        181 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLP 253

Qy        241 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 KLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 313

Qy        301 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 VHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 373

Qy        361 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKC 433

Qy        421 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVST 493

Qy        481 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 PTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 553

Qy        541 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 613

Qy        601 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||||||||||||||||||
Db        614 KEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL 658


RESULT 13
US-11-175-690-220
; Sequence 220, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 220
;   LENGTH: 655
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-220

  Query Match             95.1%;  Score 3249.5;  DB 7;  Length 655;
  Best Local Similarity   98.2%;  Pred. No. 1.1e-248;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels   11;  Gaps    1;

Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR----------DAHKSEVAHRFKDLGEEN 78
              |||||||||||||||||||||||||||||||          ||||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFDAHKSEVAHRFKDLGEEN 88

Qy         79 FKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVA 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 FKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVA 148

Qy        139 TLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKY 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 TLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKY 208

Qy        199 LYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRL 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 LYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRL 268

Qy        259 KCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRA 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 KCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRA 328

Qy        319 DLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKN 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 DLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKN 388

Qy        379 YAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFK 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 YAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFK 448

Qy        439 PLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCC 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 PLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCC 508

Qy        499 KHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYV 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 KHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYV 568
```

```
Qy        559 PKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKC 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 PKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKC 628

Qy        619 CKADDKETCFAEEGKKLVAASQAALGL 645
              ||||||||||||||||||||||||||||
Db        629 CKADDKETCFAEEGKKLVAASQAALGL 655
```

RESULT 14
US-11-175-690-225
; Sequence 225, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 225
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-225

Query Match 95.1%; Score 3249; DB 7; Length 656;
Best Local Similarity 98.1%; Pred. No. 1.2e-248;
Matches 616; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

```
Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR------------DAHKSEVAHRFKDLGEE 77
              |||||||||||||||||||||||||||||||||            ||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDAHKSEVAHRFKDLGEE 88

Qy         78 NFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTV 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 NFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTV 148

Qy        138 ATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKK 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 ATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKK 208

Qy        198 YLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQR 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 YLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQR 268

Qy        258 LKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 LKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDR 328

Qy        318 ADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCK 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCK 388

Qy        378 NYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEF 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 NYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEF 448

Qy        438 KPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKC 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 KPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKC 508

Qy        498 CKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETY 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 CKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETY 568

Qy        558 VPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 VPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK 628

Qy        618 CCKADDKETCFAEEGKKLVAASQAALGL 645
              ||||||||||||||||||||||||||||
Db        629 CCKADDKETCFAEEGKKLVAASQAALGL 656
```

RESULT 15
US-11-175-690-216
; Sequence 216, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 216
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-216

Query Match 95.1%; Score 3248.5; DB 7; Length 657;
Best Local Similarity 97.9%; Pred. No. 1.3e-248;
Matches 616; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

```
Qy         30 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR------------DAHKSEVAHRFKDLGE 76
              |||||||||||||||||||||||||||||||||            ||||||||||||||||
Db         29 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRDAHKSEVAHRFKDDAHKSEVAHRFKDLGE 88

Qy         77 ENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 ENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCT 148

Qy        137 VATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLK 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 VATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLK 208

Qy        197 KYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQ 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 KYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQ 268
```

```
Qy        257 RLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 RLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD 328

Qy        317 RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVC 388

Qy        377 KNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDE 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 KNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDE 448

Qy        437 FKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSK 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 FKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSK 508

Qy        497 CCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDET 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 CCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDET 568

Qy        557 YVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVE 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 YVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVE 628

Qy        617 KCCKADDKETCFAEEGKKLVAASQAALGL 645
              |||||||||||||||||||||||||||||
Db        629 KCCKADDKETCFAEEGKKLVAASQAALGL 657
```

Search completed: April 19, 2006, 12:36:42
Job time : 22.5305 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 11:56:31 ; Search time 21.1122 Seconds
(without alignments)
1852.232 Million cell updates/sec

Title: US-10-775-180-449
Perfect score: 465
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....SSYLEGQAAKEFIAWLVKGR 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	465	100.0	89	7	ADF16195	Adf16195 Human alb
2	465	100.0	89	7	ADH21652	Adh21652 Human GLP
3	465	100.0	673	7	ADF17044	Adf17044 Human alb
4	465	100.0	673	7	ADH22037	Adh22037 Mouse alb
5	465	100.0	674	7	ADF16193	Adf16193 Human alb
6	465	100.0	674	7	ADH21650	Adh21650 Human alb
7	465	100.0	674	9	ADW45202	Adw45202 K. lactis
8	465	100.0	915	9	ADW45204	Adw45204 K. lactis
9	341.5	73.4	669	7	ADF16144	Adf16144 Human alb
10	341.5	73.4	669	7	ADH21622	Adh21622 Human alb
11	341	73.3	145	7	ADF16688	Adf16688 Human alb
12	341	73.3	145	7	ADH21888	Adh21888 Human GLP
13	341	73.3	730	7	ADF16525	Adf16525 Human alb
14	341	73.3	730	7	ADH21813	Adh21813 Human alb
15	335.5	72.2	669	7	ADF16150	Adf16150 Human alb
16	335.5	72.2	669	7	ADH21628	Adh21628 Human alb
17	335	72.0	145	7	ADF16690	Adf16690 Human alb
18	335	72.0	145	7	ADH21890	Adh21890 Human GLP
19	335	72.0	730	7	ADF16527	Adf16527 Human alb
20	335	72.0	730	7	ADH21815	Adh21815 Human alb
21	329.5	70.9	669	7	ADF16149	Adf16149 Human alb
22	329.5	70.9	669	7	ADF16148	Adf16148 Human alb
23	329.5	70.9	669	7	ADF16145	Adf16145 Human alb
24	329.5	70.9	669	7	ADF16146	Adf16146 Human alb
25	329.5	70.9	669	7	ADH21624	Adh21624 Human alb
26	329.5	70.9	669	7	ADH21626	Adh21626 Human alb
27	329.5	70.9	669	7	ADH21623	Adh21623 Human alb
28	329.5	70.9	669	7	ADH21627	Adh21627 Human alb
29	324	69.7	83	7	ADF16687	Adf16687 Human alb
30	324	69.7	83	7	ADH21887	Adh21887 Human GLP
31	324	69.7	668	7	ADF16524	Adf16524 Human alb
32	324	69.7	668	7	ADH21812	Adh21812 Human alb
33	319	68.6	77	7	ADF16689	Adf16689 Human alb
34	319	68.6	77	7	ADH21889	Adh21889 Human GLP
35	319	68.6	662	7	ADF16526	Adf16526 Human alb
36	319	68.6	662	7	ADH21814	Adh21814 Human alb
37	318	68.4	83	7	ADF16691	Adf16691 Human alb
38	318	68.4	83	7	ADH21891	Adh21891 Human GLP
39	318	68.4	668	7	ADF16528	Adf16528 Human alb
40	318	68.4	668	7	ADH21816	Adh21816 Human alb
41	317.5	68.3	664	7	ADF16510	Adf16510 Human alb
42	317.5	68.3	664	7	ADH21801	Adh21801 Human alb
43	317	68.2	647	9	ADW45208	Adw45208 K. lactis
44	315.5	67.8	663	7	ADF16512	Adf16512 Human alb
45	315.5	67.8	663	7	ADH21803	Adh21803 Human alb

ALIGNMENTS

RESULT 1
ADF16195
ID ADF16195 standard; protein; 89 AA.
XX
AC ADF16195;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human albumin fusion protein-related protein SeqID1282.
XX
KW albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human.
XX
OS Homo sapiens.
XX
PN WO2003060071-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040891.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
PA (PRIN-) PRINCIPIA PHARM CORP.

XX
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598517/56.
DR N-PSDB; ADF16194.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Example 4; SEQ ID NO 1282; 24pp; English.
XX
CC This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is that of a therapeutic protein
CC which was fused with human albumin to create a novel albumin fusion
CC protein of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ Sequence 89 AA;

Query Match 100.0%; Score 465; DB 7; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e-45;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
Db 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy 61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
Db 61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

RESULT 2
ADH21652
ID ADH21652 standard; protein; 89 AA.
XX
AC ADH21652;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human GLP-1(7-36(A8G))x2, SEQ ID NO:449.
XX
KW Fusion protein; human serum albumin; HSA; therapeutic protein;
KW shelf-life; in vitro biological activity; in vivo biological activity;
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW anorectic; ophthalmological; gene therapy; mutant; mutein.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO2003059934-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040892.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598501/56.
DR N-PSDB; ADH21651.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Disclosure; SEQ ID NO 449; 1086pp; English.
XX
CC The invention relates to fusion proteins comprising human serum albumin
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC antibody or peptide or their variants or fragments. The therapeutic
CC protein may be fused to the N-terminus, the C-terminus or both termini of
CC albumin via a linker. The albumin component of the fusion proteins
CC prolongs the shelf-life and the in vitro and vivo biological activity of
CC the proteins compared with those of the corresponding therapeutic
CC proteins on their own. The invention also relates to nucleic acids
CC encoding albumin fusion proteins, vectors and host cells comprising an
CC albumin fusion protein nucleic acid, compositions and kits comprising an
CC albumin fusion protein, the method of extending the shelf-life of a
CC therapeutic protein by fusion with albumin, and the treatment of disease
CC using an albumin fusion protein. The albumin fusion proteins may be used
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC related conditions. Specifically the albumin fusion proteins may be used
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC (especially neuropathy), retinopathy, cardiovascular disorders
CC (especially heart disease, renal disorders and obesity. The proteins may
CC also be used in a method of maintaining a basal glucose level in a
CC patient and in a method for losing weight. The present sequence is
CC related to the invention.
XX
SQ Sequence 89 AA;

Query Match 100.0%; Score 465; DB 7; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.3e-45;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
Db 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy 61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
Db 61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

RESULT 3
ADF17044
ID ADF17044 standard; protein; 673 AA.
XX
AC ADF17044;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human albumin therapeutic fusion protein SeqID2170.
XX
KW albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human.

XX
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003060071-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040891.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
PA (PRIN-) PRINCIPIA PHARM CORP.
XX
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598517/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Example 4; SEQ ID NO 2170; 24pp; English.
XX
CC This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is the amino acid sequence of a
CC novel full-length human albumin therapeutic fusion protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ Sequence 673 AA;

Query Match 100.0%; Score 465; DB 7; Length 673;
Best Local Similarity 100.0%; Pred. No. 1.4e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy 61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
|||||||||||||||||||||||||||||
Db 61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89

RESULT 4
ADH22037
ID ADH22037 standard; protein; 673 AA.
XX
AC ADH22037;
XX
DT 11-MAR-2004 (first entry)
XX
DE Mouse albumin/human GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:834.
XX
KW Fusion protein; human serum albumin; HSA; therapeutic protein;
KW shelf-life; in vitro biological activity; in vivo biological activity;
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW anorectic; ophthalmological; gene therapy; mouse serum albumin.
XX
OS Synthetic.
OS Chimeric.
OS Homo sapiens.
OS Mus sp.
XX
PN WO2003059934-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040892.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598501/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Disclosure; SEQ ID NO 834; 1086pp; English.
XX
CC The invention relates to fusion proteins comprising human serum albumin
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC antibody or peptide or their variants or fragments. The therapeutic
CC protein may be fused to the N-terminus, the C-terminus or both termini of
CC albumin via a linker. The albumin component of the fusion proteins
CC prolongs the shelf-life and the in vitro and vivo biological activity of
CC the proteins compared with those of the corresponding therapeutic
CC proteins on their own. The invention also relates to nucleic acids
CC encoding albumin fusion proteins, vectors and host cells comprising an
CC albumin fusion protein nucleic acid, compositions and kits comprising an
CC albumin fusion protein, the method of extending the shelf-life of a
CC therapeutic protein by fusion with albumin, and the treatment of disease
CC using an albumin fusion protein. The albumin fusion proteins may be used
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC related conditions. Specifically the albumin fusion proteins may be used

```
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 673 AA;

  Query Match             100.0%;  Score 465;  DB 7;  Length 673;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-44;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 5
ADF16193
ID   ADF16193 standard; protein; 674 AA.
XX
AC   ADF16193;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1280.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1280; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 674 AA;

  Query Match             100.0%;  Score 465;  DB 7;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-44;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 6
ADH21650
ID   ADH21650 standard; protein; 674 AA.
XX
AC   ADH21650;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G))x2 fusion protein, SEQ ID NO:447.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
```

PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598501/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Disclosure; SEQ ID NO 447; 1086pp; English.
XX
CC The invention relates to fusion proteins comprising human serum albumin
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC antibody or peptide or their variants or fragments. The therapeutic
CC protein may be fused to the N-terminus, the C-terminus or both termini of
CC albumin via a linker. The albumin component of the fusion proteins
CC prolongs the shelf-life and the in vitro and vivo biological activity of
CC the proteins compared with those of the corresponding therapeutic
CC proteins on their own. The invention also relates to nucleic acids
CC encoding albumin fusion proteins, vectors and host cells comprising an
CC albumin fusion protein nucleic acid, compositions and kits comprising an
CC albumin fusion protein, the method of extending the shelf-life of a
CC therapeutic protein by fusion with albumin, and the treatment of disease
CC using an albumin fusion protein. The albumin fusion proteins may be used
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC related conditions. Specifically the albumin fusion proteins may be used
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC (especially neuropathy), retinopathy, cardiovascular disorders
CC (especially heart disease, renal disorders and obesity. The proteins may
CC also be used in a method of maintaining a basal glucose level in a
CC patient and in a method for losing weight. The present sequence is
CC related to the invention.
XX
SQ Sequence 674 AA;

```
  Query Match             100.0%;  Score 465;  DB 7;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-44;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
```

RESULT 7
ADW45202
ID ADW45202 standard; protein; 674 AA.
XX
AC ADW45202;
XX
DT 07-APR-2005 (first entry)
XX
DE K. lactis killer toxin-GLP1-human serum albumin fusion protein - SEQ 206.
XX
KW fusion protein; anti-HIV; gastrointestinal-gen.; antidiabetic; anorectic;
KW nephrotropic; cardiant; cytostatic; neuroprotective; immunosuppressive;
KW immune disorder; hematological disease; hyperproliferative disorder;
KW renal disease; cardiovascular disease; cardiovascular-gen.;
KW respiratory disorder; angiogenesis disorder; neurological disease;
KW wound healing; vulnerary; endocrine disease; reproductive disorder;
KW gynecological; infectious disease; antimicrobial;
KW gastrointestinal disease; gene therapy; toxin; HSA; albumin;
KW glucagon-like peptide 1; GLP1.
XX
OS Homo sapiens.
OS Kluyveromyces lactis.
OS Chimeric.
XX
PN WO2005003296-A2.
XX
PD 13-JAN-2005.
XX
PF 20-JAN-2004; 2004WO-US001369.
XX
PR 22-JAN-2003; 2003US-0441305P.
PR 11-MAR-2003; 2003US-0453201P.
PR 02-MAY-2003; 2003US-0467222P.
PR 23-MAY-2003; 2003US-0472816P.
PR 06-JUN-2003; 2003US-0476267P.
PR 24-SEP-2003; 2003US-0505172P.
PR 30-SEP-2003; 2003US-0506746P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Haseltine WA, Rosen CA;
XX
DR WPI; 2005-091786/10.
XX
PT New albumin fusion protein for diagnosing, treating or preventing
PT diseases such as HIV/AIDS, diabetes, obesity, heart disease or immune
PT disorders comprises a therapeutic protein (e.g. CD4M33, GLP-2 or PACAP-
PT 27) and an albumin.
XX
PS Example 13; SEQ ID NO 206; 884pp; English.
XX
CC The invention relates to a novel albumin fusion protein comprising a
CC therapeutic protein as listed in the specification in Table 1 and an
CC albumin comprising a sequence of SEQ ID NO: 1, or a fragment or variant
CC of SEQ ID NO: 1, where the fragment or variant has albumin activity and
CC where the albumin activity is the ability to prolong the shelf life of
CC the therapeutic protein compared to the shelf-life of the therapeutic
CC protein in an unfused state. Human serum albumin (HSA, HA) is responsible
CC for a significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion
CC protein of the invention demonstrates anti-HIV, gastrointestinal-gen.,
CC antidiabetic, anorectic, cardiant and immunosuppressive activities. The
CC fusion protein may be useful for diagnosing, treating, preventing or
CC ameliorating diseases, such as immune disorders, blood disorders,
CC hyperproliferative disorders, renal disorders, cardiovascular disorders,
CC respiratory disorders, angiogenesis-related disorders, neurological
CC disorders, wound healing disorders, endocrine disorders, reproductive
CC disorders, infectious disorders and gastrointestinal disorders, possibly
CC with the use of gene therapy techniques. The current sequence is that of
CC the Kluyveromyces lactis killer toxin-GLP1-human serum albumin fusion
CC protein - SEQ 206 of the invention.
XX
SQ Sequence 674 AA;

```
  Query Match             100.0%;  Score 465;  DB 9;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-44;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
```

RESULT 8
ADW45204
ID ADW45204 standard; protein; 915 AA.

```
XX
AC   ADW45204;
XX
DT   07-APR-2005  (first entry)
XX
DE   K. lactis killer toxin-GLP1-HSA-GFP tag fusion protein - SEQ ID 208.
XX
KW   fusion protein; anti-HIV; gastrointestinal-gen.; antidiabetic; anorectic;
KW   nephrotropic; cardiant; cytostatic; neuroprotective; immunosuppressive;
KW   immune disorder; hematological disease; hyperproliferative disorder;
KW   renal disease; cardiovascular disease; cardiovascular-gen.;
KW   respiratory disorder; angiogenesis disorder; neurological disease;
KW   wound healing; vulnerary; endocrine disease; reproductive disorder;
KW   gynecological; infectious disease; antimicrobial;
KW   gastrointestinal disease; gene therapy; toxin; HSA; albumin;
KW   glucagon-like peptide 1; GLP1.
XX
OS   Homo sapiens.
OS   Kluyveromyces lactis.
OS   Chimeric.
XX
PN   WO2005003296-A2.
XX
PD   13-JAN-2005.
XX
PF   20-JAN-2004; 2004WO-US001369.
XX
PR   22-JAN-2003; 2003US-0441305P.
PR   11-MAR-2003; 2003US-0453201P.
PR   02-MAY-2003; 2003US-0467222P.
PR   23-MAY-2003; 2003US-0472816P.
PR   06-JUN-2003; 2003US-0476267P.
PR   24-SEP-2003; 2003US-0505172P.
PR   30-SEP-2003; 2003US-0506746P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Haseltine WA,  Rosen CA;
XX
DR   WPI; 2005-091786/10.
XX
PT   New albumin fusion protein for diagnosing, treating or preventing
PT   diseases such as HIV/AIDS, diabetes, obesity, heart disease or immune
PT   disorders comprises a therapeutic protein (e.g. CD4M33, GLP-2 or PACAP-
PT   27) and an albumin.
XX
PS   Example 13; SEQ ID NO 208; 884pp; English.
XX
CC   The invention relates to a novel albumin fusion protein comprising a
CC   therapeutic protein as listed in the specification in Table 1 and an
CC   albumin comprising a sequence of SEQ ID NO: 1, or a fragment or variant
CC   of SEQ ID NO: 1, where the fragment or variant has albumin activity and
CC   where the albumin activity is the ability to prolong the shelf life of
CC   the therapeutic protein compared to the shelf-life of the therapeutic
CC   protein in an unfused state. Human serum albumin (HSA, HA) is responsible
CC   for a significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion
CC   protein of the invention demonstrates anti-HIV, gastrointestinal-gen.,
CC   antidiabetic, anorectic, cardiant and immunosuppressive activities. The
CC   fusion protein may be useful for diagnosing, treating, preventing or
CC   ameliorating diseases, such as immune disorders, blood disorders,
CC   hyperproliferative disorders, renal disorders, cardiovascular disorders,
CC   respiratory disorders, angiogenesis-related disorders, neurological
CC   disorders, wound healing disorders, endocrine disorders, reproductive
CC   disorders, infectious disorders and gastrointestinal disorders, possibly
CC   with the use of gene therapy techniques. The current sequence is that of
CC   the Kluyveromyces lactis killer toxin-GLP1-HSA-GFP tag fusion protein -
CC   SEQ ID 208 of the invention.
XX
SQ   Sequence 915 AA;

  Query Match             100.0%;  Score 465;  DB 9;  Length 915;
  Best Local Similarity   100.0%;  Pred. No. 2e-44;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 9
ADF16144
ID   ADF16144 standard; protein; 669 AA.
XX
AC   ADF16144;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1231.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1231; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
```

```
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 669 AA;

  Query Match             73.4%;  Score 341.5;  DB 7;  Length 669;
  Best Local Similarity   79.3%;  Pred. No. 2.2e-30;
  Matches   69;  Conservative    4;  Mismatches    5;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||||||||||||||||||||||||||||||||||
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGR 84


RESULT 10
ADH21622
ID   ADH21622 standard; protein; 669 AA.
XX
AC   ADH21622;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:419.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Disclosure; SEQ ID NO 419; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 669 AA;

  Query Match             73.4%;  Score 341.5;  DB 7;  Length 669;
  Best Local Similarity   79.3%;  Pred. No. 2.2e-30;
  Matches   69;  Conservative    4;  Mismatches    5;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||||||||||||||||||||||||||||||||||
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGR 84


RESULT 11
ADF16688
ID   ADF16688 standard; protein; 145 AA.
XX
AC   ADF16688;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin fusion protein-related protein SeqID1790.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
```

```
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
DR   N-PSDB; ADF16362.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1790; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is that of a therapeutic protein
CC   which was fused with human albumin to create a novel albumin fusion
CC   protein of the invention. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 145 AA;

  Query Match             73.3%;  Score 341;  DB 7;  Length 145;
  Best Local Similarity   81.7%;  Pred. No. 4.3e-31;
  Matches   67;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGR 145


RESULT 12
ADH21888
ID   ADH21888 standard; protein; 145 AA.
XX
AC   ADH21888;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human GLP-1(7-36(A8G)), SEQ ID NO:685.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy; mutant; mutein.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
DR   N-PSDB; ADH21736.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Disclosure; SEQ ID NO 685; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 145 AA;

  Query Match             73.3%;  Score 341;  DB 7;  Length 145;
  Best Local Similarity   81.7%;  Pred. No. 4.3e-31;
  Matches   67;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123
```

```
Qy          68 DVSSYLEGQAAKEFIAWLVKGR 89
               ||||||||||||||||||||||
Db         124 DVSSYLEGQAAKEFIAWLVKGR 145


RESULT 13
ADF16525
ID   ADF16525 standard; protein; 730 AA.
XX
AC   ADF16525;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1622.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1622; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 730 AA;

  Query Match             73.3%;  Score 341;  DB 7;  Length 730;
  Best Local Similarity   81.7%;  Pred. No. 2.8e-30;
  Matches   67;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy           8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
               ||: : :  :       ||||||||||||||||||||||||||||||||||||||||||||
Db          64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy          68 DVSSYLEGQAAKEFIAWLVKGR 89
               ||||||||||||||||||||||
Db         124 DVSSYLEGQAAKEFIAWLVKGR 145


RESULT 14
ADH21813
ID   ADH21813 standard; protein; 730 AA.
XX
AC   ADH21813;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human albumin/GLP-1(7-36(A8G)) fusion protein, SEQ ID NO:610.
XX
KW   Fusion protein; human serum albumin; HSA; therapeutic protein;
KW   shelf-life; in vitro biological activity; in vivo biological activity;
KW   metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW   diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW   retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW   obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW   anorectic; ophthalmological; gene therapy.
XX
OS   Synthetic.
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003059934-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040892.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598501/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
```

```
XX
PS   Disclosure; SEQ ID NO 610; 1086pp; English.
XX
CC   The invention relates to fusion proteins comprising human serum albumin
CC   (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC   antibody or peptide or their variants or fragments. The therapeutic
CC   protein may be fused to the N-terminus, the C-terminus or both termini of
CC   albumin via a linker. The albumin component of the fusion proteins
CC   prolongs the shelf-life and the in vitro and vivo biological activity of
CC   the proteins compared with those of the corresponding therapeutic
CC   proteins on their own. The invention also relates to nucleic acids
CC   encoding albumin fusion proteins, vectors and host cells comprising an
CC   albumin fusion protein nucleic acid, compositions and kits comprising an
CC   albumin fusion protein, the method of extending the shelf-life of a
CC   therapeutic protein by fusion with albumin, and the treatment of disease
CC   using an albumin fusion protein. The albumin fusion proteins may be used
CC   in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC   related conditions. Specifically the albumin fusion proteins may be used
CC   to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC   (especially neuropathy), retinopathy, cardiovascular disorders
CC   (especially heart disease, renal disorders and obesity. The proteins may
CC   also be used in a method of maintaining a basal glucose level in a
CC   patient and in a method for losing weight. The present sequence is
CC   related to the invention.
XX
SQ   Sequence 730 AA;

  Query Match             73.3%;  Score 341;  DB 7;  Length 730;
  Best Local Similarity   81.7%;  Pred. No. 2.8e-30;
  Matches   67;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGR 145


RESULT 15
ADF16150
ID   ADF16150 standard; protein; 669 AA.
XX
AC   ADF16150;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human albumin therapeutic fusion protein SeqID1237.
XX
KW   albumin fusion protein; albumin activity; human serum albumin;
KW   serum osmotic pressure; shelf-life; stability; antidiabetic;
KW   gene therapy; diabetes mellitus; human.
XX
OS   Chimeric.
OS   Homo sapiens.
XX
PN   WO2003060071-A2.
XX
PD   24-JUL-2003.
XX
PF   23-DEC-2002; 2002WO-US040891.
XX
PR   21-DEC-2001; 2001US-0341811P.
PR   24-JAN-2002; 2002US-0350358P.
PR   28-JAN-2002; 2002US-0351360P.
PR   26-FEB-2002; 2002US-0359370P.
PR   28-FEB-2002; 2002US-0360000P.
PR   27-MAR-2002; 2002US-0367500P.
PR   08-APR-2002; 2002US-0370227P.
PR   10-MAY-2002; 2002US-0378950P.
PR   24-MAY-2002; 2002US-0382617P.
PR   28-MAY-2002; 2002US-0383123P.
PR   05-JUN-2002; 2002US-0385708P.
PR   10-JUL-2002; 2002US-0394625P.
PR   24-JUL-2002; 2002US-0398008P.
PR   09-AUG-2002; 2002US-0402131P.
PR   13-AUG-2002; 2002US-0402708P.
PR   18-SEP-2002; 2002US-0411355P.
PR   18-SEP-2002; 2002US-0411426P.
PR   02-OCT-2002; 2002US-0414984P.
PR   11-OCT-2002; 2002US-0417611P.
PR   23-OCT-2002; 2002US-0420246P.
PR   05-NOV-2002; 2002US-0423623P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA   (PRIN-) PRINCIPIA PHARM CORP.
XX
PI   Ballance DJ,  Turner AJ,  Rosen CA,  Haseltine WA;
XX
DR   WPI; 2003-598517/56.
XX
PT   New albumin fusion protein, useful for preparing a composition for
PT   treating diabetes mellitus.
XX
PS   Example 4; SEQ ID NO 1237; 24pp; English.
XX
CC   This invention relates to a novel albumin fusion protein having albumin
CC   or biological activity. Human serum albumin is responsible for a
CC   significant proportion of the osmotic pressure of serum and also
CC   functions as a carrier of endogenous and exogenous ligands. The fusion of
CC   albumin to a therapeutic protein may increase shelf-life and stability of
CC   the therapeutic protein. The albumin fusion protein of the invention may
CC   allow production of compositions with antidiabetic activity whilst the
CC   nucleotide sequence which encodes it may be useful for gene therapy. The
CC   albumin fusion protein is useful for preparing a composition for treating
CC   diabetes mellitus. The present sequence is the amino acid sequence of a
CC   novel full-length human albumin therapeutic fusion protein of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ   Sequence 669 AA;

  Query Match             72.2%;  Score 335.5;  DB 7;  Length 669;
  Best Local Similarity   78.2%;  Pred. No. 1.1e-29;
  Matches   68;  Conservative    4;  Mismatches    6;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :          |||||||||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGR 84


Search completed: April 19, 2006, 12:02:33
Job time : 21.1122 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:02:52 ; Search time 4.04545 Seconds
(without alignments)
2116.769 Million cell updates/sec

Title: US-10-775-180-449
Perfect score: 465
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....SSYLEGQAAKEFIAWLVKGR 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	247.5	53.2	180	1	GCBO	glucagon precursor
2	243.5	52.4	180	1	GCHU	glucagon precursor
3	243.5	52.4	180	2	A57294	glucagon precursor
4	241.5	51.9	180	1	GCHY	glucagon precursor
5	241.5	51.9	180	1	GCRT	glucagon precursor
6	240.5	51.7	180	1	GCGP	glucagon precursor
7	238.5	51.3	158	1	GCPG	glucagon precursor
8	230.5	49.6	180	1	GCRTDU	glucagon precursor
9	225.5	48.5	151	1	GCCH	glucagon precursor
10	225.5	48.5	206	2	I51301	proglucagon - chic
11	208.5	44.8	178	2	I51058	glucagon I precurs
12	206	44.3	101	1	GCFGB	glucagon precursor
13	200.5	43.1	122	1	GCAF2	glucagon 2 precurs
14	200.5	43.1	178	2	I51057	glucagon II precur
15	188.5	40.5	63	1	GCIDC	glucagon precursor
16	183	39.4	124	1	GCAF	glucagon 1 precurs
17	182	39.1	72	1	GCGXA	glucagon precursor
18	179.5	38.6	87	1	GCFIS	glucagon precursor
19	178.5	38.4	60	1	GCONC	glucagon precursor
20	155	33.3	66	2	I51093	glucagon - chinook
21	142	30.5	1146	2	S07915	RP2 protein - yeas
22	122	26.2	30	2	C61125	glucagon-like pept
23	122	26.2	30	2	B61125	glucagon-like pept
24	107	23.0	30	2	S44473	glucagon-like pept
25	106	22.8	69	1	GCDG69	glucagon-69 - dog
26	96	20.6	29	2	S07211	glucagon - marbled
27	95	20.4	31	2	S44472	glucagon G2 - Nort
28	95	20.4	39	1	HWGH4G	exendin-4 - Gila m
29	94	20.2	29	1	GCDF	glucagon - smaller
30	93	20.0	31	2	S44471	glucagon G1 - Nort
31	92	19.8	29	1	GCEN	glucagon - elephan
32	89	19.1	29	1	GCOPV	glucagon - North A
33	89	19.1	29	2	A91740	glucagon - turkey
34	89	19.1	29	2	C39258	glucagon - common
35	89	19.1	29	2	A91742	glucagon - Arabian
36	89	19.1	29	2	A91741	glucagon - rabbit
37	87	18.7	29	1	A61583	glucagon - ostrich
38	87	18.7	29	1	GCDK	glucagon - duck
39	87	18.7	29	1	GCTTS	glucagon - slider
40	87	18.7	29	2	C60840	glucagon I - Europ
41	86.5	18.6	290	2	S52860	neuropeptide Pol-R
42	86	18.5	29	1	GCCB	glucagon - Chinchi
43	86	18.5	39	1	HWGH3Z	exendin-3 - Mexica
44	85	18.3	29	1	GCFLE	glucagon - Europea
45	85	18.3	29	2	A61135	glucagon - bigeye

ALIGNMENTS

RESULT 1
GCBO
glucagon precursor - bovine
N;Contains: glicentin-related peptide; glucagon; glucagon-like peptide 1; glucagon-like
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 20-Mar-1998
C;Accession: A93970; A92081; A01538
R;Lopez, L.C.; Frazier, M.L.; Su, C.J.; Kumar, A.; Saunders, G.F.
Proc. Natl. Acad. Sci. U.S.A. 80, 5485-5489, 1983
A;Title: Mammalian pancreatic preproglucagon contains three glucagon-related peptides.
A;Reference number: A93970; MUID:83299996; PMID:6577439
A;Accession: A93970
A;Molecule type: mRNA
A;Residues: 1-180 <LOP>
A;Cross-references: UNIPARC:UPI00001734FF; EMBL:K00107
R;Bromer, W.W.; Boucher, M.E.; Koffenberger Jr., J.E.
J. Biol. Chem. 246, 2822-2827, 1971
A;Title: Amino acid sequence of bovine glucagon.
A;Reference number: A92081; MUID:71166445; PMID:5102927
A;Accession: A92081
A;Molecule type: protein
A;Residues: 53-81 <BRO>
A;Cross-references: UNIPARC:UPI000002C586
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pancr
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-81/Product: glucagon #status experimental <GCN>
F;98-127/Product: glucagon-like peptide 1 #status experimental <GL1>
F;146-178/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following g

Query Match 53.2%; Score 247.5; DB 1; Length 180;
Best Local Similarity 42.4%; Pred. No. 3e-18;
Matches 53; Conservative 16; Mismatches 19; Indels 37; Gaps 3;

```
Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRGSL----------------DKRHGEGTFTSDV 39
              :::::    :  |||     |::|  : |                  |||| :||||||
Db          3 SLYFVAGLFVMLVQGSWQRSLQNTEEKSSSFPAPQTDPLGDPDQINEDKRHSQGTFTSDY 62

Qy         40 SSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSSYLEGQAAKEFIAW 84
               | ||: : |::|: ||:               || ||||||||||||||||||||||
Db         63 SKYLDSRRAQDFVQWLMNTKRNKNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAW 122

Qy         85 LVKGR 89
              |||||
Db        123 LVKGR 127
```

RESULT 2
GCHU

```
glucagon precursor [validated] - human
N;Contains: glicentin; glicentin-related polypeptide (GRPP); glucagon; glucagon-like pep
ke peptide 1 (tGLP1)
C;Species: Homo sapiens (man)
C;Date: 24-Apr-1984 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: A24377; A44197; A30875; A32614; A01541; S23309
R;White, J.W.; Saunders, G.F.
Nucleic Acids Res. 14, 4719-4730, 1986
A;Title: Structure of the human glucagon gene.
A;Reference number: A24377; MUID:86259053; PMID:3725587
A;Accession: A24377
A;Molecule type: DNA
A;Residues: 1-180 <WHI>
A;Cross-references: UNIPROT:P01275; UNIPARC:UPI000012B832; GB:X03991
R;Bell, G.I.; Sanchez-Pescador, R.; Laybourn, P.J.; Najarian, R.C.
Nature 304, 368-371, 1983
A;Title: Exon duplication and divergence in the human preproglucagon gene.
A;Reference number: A44197; MUID:83271477; PMID:6877358
A;Accession: A44197
A;Molecule type: DNA
A;Residues: 1-179 <BEL>
A;Cross-references: UNIPARC:UPI000016A9A7; GB:V01515; NID:g31777; PIDN:CAA24759.1; PID:g
R;Drucker, D.J.; Asa, S.
J. Biol. Chem. 263, 13475-13478, 1988
A;Title: Glucagon gene expression in vertebrate brain.
A;Reference number: A30875; MUID:88330860; PMID:2901414
A;Accession: A30875
A;Molecule type: mRNA
A;Residues: 1-180 <DRU>
A;Cross-references: UNIPARC:UPI000012B832; GB:J04040; NID:g183269; PIDN:AAA52567.1; PID:
R;Orskov, C.; Bersani, M.; Johnsen, A.H.; Hojrup, P.; Holst, J.J.
J. Biol. Chem. 264, 12826-12829, 1989
A;Title: Complete sequences of glucagon-like peptide-1 from human and pig small intestin
A;Reference number: A92732; MUID:89327238; PMID:2753890
A;Accession: A32614
A;Molecule type: protein
A;Residues: 98-127 <ORS>
A;Cross-references: UNIPARC:UPI0000032E2A
R;Thomsen, J.; Kristiansen, K.; Brunfeldt, K.; Sundby, F.
FEBS Lett. 21, 315-319, 1972
A;Title: The amino acid sequence of human glucagon.
A;Reference number: A91373
A;Accession: A01541
A;Molecule type: protein
A;Residues: 53-81 <THO>
A;Cross-references: UNIPARC:UPI000002C586
R;Tsugita, A.; Takamoto, K.; Kamo, M.; Iwadate, H.
Eur. J. Biochem. 206, 691-696, 1992
A;Title: C-terminal sequencing of protein. A novel partial acid hydrolysis and analysis
A;Reference number: S23188; MUID:92298996; PMID:1606956
A;Accession: S23309
A;Molecule type: protein
A;Residues: 53-81 <TSU>
A;Cross-references: UNIPARC:UPI000002C586
C;Comment: In pancreatic alpha-cells, proglucagon is processed to glicentin-related poly
stinal L cells, proglucagon is processed to truncated glucagon-like peptide 1, glucagon-
dulin.
C;Genetics:
A;Gene: GDB:GCG
A;Cross-references: GDB:119265; OMIM:138030
A;Map position: 2q36-2q37
A;Introns: 31/2; 85/2; 131/2; 179/2
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; intest
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status experimental <PGC>
F;21-89/Product: glicentin #status experimental <GLN>
F;21-50/Product: glicentin-related polypeptide #status predicted <GRPP>
F;53-89/Product: oxyntomodulin #status experimental <OXN>
F;53-81/Product: glucagon #status experimental <GCN>
F;92-178/Product: major proglucagon fragment #status experimental <MPGF>
F;92-127/Product: glucagon-like peptide 1 #status experimental <GL1>
F;98-127/Product: truncated glucagon-like peptide 1 #status experimental <TGL>
F;146-178/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             52.4%;  Score 243.5;  DB 1;  Length 180;
  Best Local Similarity   42.4%;  Pred. No. 7.6e-18;
  Matches   53;  Conservative   15;  Mismatches   20;  Indels   37;  Gaps    3;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRG----------------SLDKRHGEGTFTSDV 39
              :|:::    :  |||     |: |  :                : |||| :||||||
Db          3 SIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDPDQMNEDKRHSQGTFTSDY 62

Qy         40 SSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSSYLEGQAAKEFIAW 84
              | ||: : |::|: ||:                || ||||||||||||||||||||||
Db         63 SKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAW 122

Qy         85 LVKGR 89
              |||||
Db        123 LVKGR 127


RESULT 3
A57294
glucagon precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C;Accession: A57294; S49903
R;Rothenberg, M.E.; Eilertson, C.D.; Klein, K.; Zhou, Y.; Lindberg, I.; McDonald, J.K.;
J. Biol. Chem. 270, 10136-10146, 1995
A;Title: Processing of mouse proglucagon by recombinant prohormone convertase 1 and immu
A;Reference number: A57294; MUID:95247722; PMID:7730317
A;Accession: A57294
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-180 <ROT>
A;Cross-references: UNIPROT:P55095; UNIPARC:UPI000000192D; EMBL:Z46845; NID:g599880; PID
C;Superfamily: glucagon
C;Keywords: carbohydrate metabolism; duplication; hormone; pancreas

  Query Match             52.4%;  Score 243.5;  DB 2;  Length 180;
  Best Local Similarity   43.5%;  Pred. No. 7.6e-18;
  Matches   54;  Conservative   13;  Mismatches   20;  Indels   37;  Gaps    3;

Qy          3 IFYIFLFLLSFVQG-----LEHTHRRG----------------SLDKRHGEGTFTSDVS 40
              |:::   |:  |||     |: |                  : |||| :|||||| |
Db          4 IYFVAGLLIMLVQGSWQHALQDTEENPRSFPASQTEAHEDPDEMNEDKRHSQGTFTSDYS 63

Qy         41 SYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSSYLEGQAAKEFIAWL 85
               ||: : |::|: ||:                || ||||||||||||||||||||||
Db         64 KYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAWL 123

Qy         86 VKGR 89
              ||||
Db        124 VKGR 127


RESULT 4
GCHY
glucagon precursor - golden hamster
N;Contains: glicentin-related peptide; glucagon; glucagon-like peptide 1; glucagon-like
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 20-Mar-1998
C;Accession: A01539
R;Bell, G.I.; Santerre, R.F.; Mullenbach, G.T.
Nature 302, 716-718, 1983
A;Title: Hamster preproglucagon contains the sequence of glucagon and two related peptic
A;Reference number: A01539; MUID:83167563; PMID:6835407
A;Accession: A01539
A;Molecule type: mRNA
A;Residues: 1-180 <BEL>
A;Cross-references: UNIPARC:UPI00001734FE; EMBL:J00059
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pancre
F;1-20/Domain: signal sequence #status predicted <SIG>
```

```
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-81/Product: glucagon #status predicted <GCN>
F;98-127/Product: glucagon-like peptide 1 #status predicted <GL1>
F;146-180/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             51.9%;  Score 241.5;  DB 1;  Length 180;
  Best Local Similarity   60.3%;  Pred. No. 1.2e-17;
  Matches   47;  Conservative    7;  Mismatches    9;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSS 71
              |||| :|||||| | ||: : |::|: ||:                || ||||||||||
Db         50 DKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


RESULT 5
GCRT
glucagon precursor - rat
N;Contains: glicentin-related peptide; glucagon; glucagon-like peptide 1; glucagon-like
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C;Accession: A22655; A25190; A44198
R;Heinrich, G.; Gros, P.; Habener, J.F.
J. Biol. Chem. 259, 14082-14087, 1984
A;Title: Glucagon gene sequence: four of six exons encode separate functional domains of
A;Reference number: A22655; MUID:85054853; PMID:6094539
A;Accession: A22655
A;Molecule type: DNA
A;Residues: 1-180 <HEI>
A;Cross-references: UNIPROT:P06883; UNIPARC:UPI000002DB13; EMBL:K02809
A;Note: the authors translated the codon TTG for residue 10 as Glu and ACC for residue 5
R;Mojsov, S.; Heinrich, G.; Wilson, I.B.; Ravazzola, M.; Orci, L.; Habener, J.F.
J. Biol. Chem. 261, 11880-11889, 1986
A;Title: Preproglucagon gene expression in pancreas and intestine diversifies at the lev
A;Reference number: A25190; MUID:86304324; PMID:3528148
A;Accession: A25190
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-180 <MOJ>
A;Cross-references: UNIPARC:UPI000002DB13
R;Heinrich, G.; Gros, P.; Lund, P.K.; Bentley, R.C.; Habener, J.F.
Endocrinology 115, 2176-2181, 1984
A;Title: Pre-proglucagon messenger ribonucleic acid: nucleotide and encoded amino acid s
A;Reference number: A44198; MUID:85051023; PMID:6548696
A;Accession: A44198
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-180 <HE2>
A;Cross-references: UNIPARC:UPI000002DB13; GB:K02809; GB:K02810; GB:K02811; GB:K02812
C;Genetics:
A;Introns: 31/2; 85/2; 131/2; 179/2
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pancre
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-81/Product: glucagon #status predicted <GCN>
F;98-127/Product: glucagon-like peptide 1 #status predicted <GL1>
F;146-180/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             51.9%;  Score 241.5;  DB 1;  Length 180;
  Best Local Similarity   60.3%;  Pred. No. 1.2e-17;
  Matches   47;  Conservative    7;  Mismatches    9;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSS 71
              |||| :|||||| | ||: : |::|: ||:                || ||||||||||
Db         50 DKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


RESULT 6
GCGP
glucagon precursor - guinea pig
N;Alternate names: oxyntomodulin
N;Contains: glicentin-related peptide; glucagon; glucagon-37 (oxyntomodulin); glucagon-1
C;Species: Cavia porcellus (guinea pig)
C;Date: 30-Sep-1987 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A24856; A23849; A60323
R;Seino, S.; Welsh, M.; Bell, G.I.; Chan, S.J.; Steiner, D.F.
FEBS Lett. 203, 25-30, 1986
A;Title: Mutations in the guinea pig preproglucagon gene are restricted to a specific pc
A;Reference number: A24856; MUID:86248118; PMID:3755107
A;Accession: A24856
A;Molecule type: mRNA
A;Residues: 1-180 <SEI>
A;Cross-references: UNIPROT:P05110; UNIPARC:UPI000012B82C; DDBJ:D00014; GB:N00014; NID:g
R;Huang, C.G.; Eng, J.; Pan, Y.C.E.; Hulmes, J.D.; Yalow, R.S.
Diabetes 35, 508-512, 1986
A;Title: Guinea pig glucagon differs from other mammalian glucagons.
A;Reference number: A23849; MUID:86165412; PMID:3956884
A;Accession: A23849
A;Molecule type: protein
A;Residues: 53-81 <HUA>
A;Cross-references: UNIPARC:UPI00001734FD
R;Conlon, J.M.; Hansen, H.F.; Schwartz, T.W.
Regul. Pept. 11, 309-320, 1985
A;Title: Primary structure of glucagon and a partial sequence of oxyntomodulin (glucagor
A;Reference number: A60323; MUID:86017849; PMID:4048553
A;Accession: A60323
A;Molecule type: protein
A;Residues: 53-81 <CON>
A;Cross-references: UNIPARC:UPI00001734FD
A;Note: glucagon-37 was not completely sequenced
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pancr
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-89/Product: glucagon-37 (oxyntomodulin) #status experimental <G37>
F;53-81/Product: glucagon #status experimental <GCN>
F;98-127/Product: glucagon-like peptide 1 #status predicted <GL1>
F;146-178/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             51.7%;  Score 240.5;  DB 1;  Length 180;
  Best Local Similarity   60.3%;  Pred. No. 1.5e-17;
  Matches   47;  Conservative    7;  Mismatches    9;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSS 71
              |||| :|||||| | ||: : |::|: ||:                || ||||||||||
Db         50 DKRHSQGTFTSDYSKYLDSRRAQQFLKWLLNVKRNRNNIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


RESULT 7
GCPG
glucagon precursor - pig (fragment)
N;Alternate names: glicentin; oxyntomodulin
N;Contains: glicentin-related peptide; glucagon; glucagon-37 (oxyntomodulin); glucagon-(
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 17-Dec-1982 #sequence_revision 31-Mar-1993 #text_change 20-Mar-1998
C;Accession: A01540; A60312; A91781; B32614; A28064
R;Thim, L.; Moody, A.J.
Regul. Pept. 2, 139-150, 1981
```

```
A;Title: The primary structure of porcine glicentin (proglucagon).
A;Reference number: A94233; MUID:81248172; PMID:6894800
A;Accession: A01540
A;Molecule type: protein
A;Residues: 1-69 <TH1>
A;Cross-references: UNIPARC:UPI0000173500
R;Thim, L.; Moody, A.J.
Regul. Pept. Suppl. 2, S33, 1983
A;Title: Primary structure of a possible porcine proglucagon fragment.
A;Reference number: A60312
A;Accession: A60312
A;Molecule type: protein
A;Residues: 1-30 <TH2>
A;Cross-references: UNIPARC:UPI000002C9AC
A;Note: this peptide is co-secreted with glucagon from the pancreas
R;Bromer, W.W.; Sinn, L.G.; Behrens, O.K.
J. Am. Chem. Soc. 79, 2807-2810, 1957
A;Title: The amino acid sequence of glucagon. V. Location of amide groups, acid degradat
A;Reference number: A91781
A;Accession: A91781
A;Molecule type: protein
A;Residues: 33-61 <BRO>
A;Cross-references: UNIPARC:UPI000002C586
R;Orskov, C.; Bersani, M.; Johnsen, A.H.; Hojrup, P.; Holst, J.J.
J. Biol. Chem. 264, 12826-12829, 1989
A;Title: Complete sequences of glucagon-like peptide-1 from human and pig small intestin
A;Reference number: A92732; MUID:89327238; PMID:2753890
A;Accession: B32614
A;Molecule type: protein
A;Residues: 78-107 <ORS>
A;Cross-references: UNIPARC:UPI0000032E2A
R;Buhl, T.; Thim, L.; Kofod, H.; Orskov, C.; Harling, H.; Holst, J.J.
J. Biol. Chem. 263, 8621-8624, 1988
A;Title: Naturally occurring products of proglucagon 111-160 in the porcine and human sm
A;Reference number: A28064; MUID:88243712; PMID:3379036
A;Accession: A28064
A;Molecule type: protein
A;Residues: 111-158 <BUH>
A;Cross-references: UNIPARC:UPI0000173501
C;Comment: X's represent missing amino acids, mostly basic, that are predicted to exist
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; intest
F;1-69/Product: glucagon-69 #status experimental <G69>
F;1-30/Region: glicentin-related peptide #status experimental
F;33-69/Product: glucagon-37 #status predicted <G37>
F;33-61/Product: glucagon #status experimental <GCN>
F;78-107/Product: glucagon-like peptide 1 #status experimental <GL1>
F;126-158/Product: glucagon-like peptide 2 #status experimental <GL2>
F;107/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             51.3%;  Score 238.5;  DB 1;  Length 158;
  Best Local Similarity   59.0%;  Pred. No. 2.2e-17;
  Matches   46;  Conservative    8;  Mismatches    9;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR--------------HGEGTFTSDVSS 71
              |||| :|||||| | ||: : |::|: ||:  :              | ||||||||||
Db         30 DKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNKNNIAXXXXXXXXHAEGTFTSDVSS 89

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db         90 YLEGQAAKEFIAWLVKGR 107


RESULT 8
GCRTDU
glucagon precursor - degu
N;Contains: glicentin-related peptide; glucagon; glucagon-like peptide 1; glucagon-like
C;Species: Octodon degus (degu)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: C36118
R;Nishi, M.; Steiner, D.F.
Mol. Endocrinol. 4, 1192-1198, 1990
A;Title: Cloning of complementary DNAs encoding islet amyloid polypeptide, insulin, and
A;Reference number: A36118; MUID:91155952; PMID:2293024
A;Accession: C36118
A;Molecule type: mRNA
A;Residues: 1-180 <NIS>
A;Cross-references: UNIPROT:P22890; UNIPARC:UPI000012B839; GB:M57688; NID:g202467; PIDN:
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pancre
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-81/Product: glucagon #status predicted <GCN>
F;98-127/Product: glucagon-like peptide 1 #status predicted <GL1>
F;146-178/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             49.6%;  Score 230.5;  DB 1;  Length 180;
  Best Local Similarity   59.0%;  Pred. No. 1.6e-16;
  Matches   46;  Conservative    7;  Mismatches   10;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSS 71
              |||| :|||||| | :|: : |::|: ||                || ||||||||||
Db         50 DKRHSQGTFTSDYSKFLDTRRAQDFLDWLKNTKRNRNEIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


RESULT 9
GCCH
glucagon precursor - chicken
N;Contains: glucagon; glucagon-like peptide 1
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1991 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: S09992; A92189; A60836; A01542
R;Hasegawa, S.; Terazono, K.; Nata, K.; Takada, T.; Yamamoto, H.; Okamoto, H.
FEBS Lett. 264, 117-120, 1990
A;Title: Nucleotide sequence determination of chicken glucagon precursor cDNA. Chicken p
A;Reference number: S09992; MUID:90249492; PMID:2338135
A;Accession: S09992
A;Molecule type: mRNA
A;Residues: 1-151 <HAS>
A;Cross-references: UNIPROT:P01277; UNIPARC:UPI000002AA99; EMBL:Y07539; NID:g63749; PIDN
R;Pollock, H.G.; Kimmel, J.R.
J. Biol. Chem. 250, 9377-9380, 1975
A;Title: Chicken glucagon. Isolation and amino acid sequence studies.
A;Reference number: A92189; MUID:76069271; PMID:1194290
A;Accession: A92189
A;Molecule type: protein
A;Residues: 55-83 <POL>
A;Cross-references: UNIPARC:UPI000012B830
R;Huang, J.; Eng, J.; Yalow, R.S.
Horm. Metab. Res. 19, 542-544, 1987
A;Title: Chicken glucagon: sequence and potency in receptor assay.
A;Reference number: A60836; MUID:88113418; PMID:2828209
A;Accession: A60836
A;Molecule type: protein
A;Residues: 55-83 <HUA>
A;Cross-references: UNIPARC:UPI000012B830
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pancre
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-151/Product: proglucagon #status predicted <PGC>
F;55-83/Product: glucagon #status experimental <GCN>
F;118-147/Product: glucagon-like peptide 1 #status predicted <GL1>
F;147/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             48.5%;  Score 225.5;  DB 1;  Length 151;
  Best Local Similarity   37.1%;  Pred. No. 4.4e-16;
  Matches   53;  Conservative   15;  Mismatches   20;  Indels   55;  Gaps    4;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRG---------SLD--------KRHGEGTFTSDV 39
              :|::|   ||  |||     |: |  :           ||        ||| :||||||
```

```
Db          5 SIYFIAGLLLMIVQGSWQNPLQDTEEKSRSFKASQSEPLDESRQLNEVKRHSQGTFTSDY 64

Qy         40 SSYLEGQAAKEFIAWLVK--------------------------------GRHGEGTFT 66
              | ||: : |::|: ||:                                ||  |||:|
Db         65 SKYLDSRRAQDFVQWLMSTKRNGQQGQEDKENDKFPDQLSSNAISKRHSEFERHAEGTYT 124

Qy         67 SDVSSYLEGQAAKEFIAWLVKGR 89
              ||::|||||||||||||||| ||
Db        125 SDITSYLEGQAAKEFIAWLVNGR 147


RESULT 10
I51301
proglucagon - chicken
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51301
R;Irwin, D.M.; Wong, J.
Mol. Endocrinol. 9, 267-277, 1995
A;Title: Trout and chicken proglucagon: alternative splicing generates mRNA transcripts
A;Reference number: A55895; MUID:95295739; PMID:7776976
A;Accession: I51301
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-206 <IRW>
A;Cross-references: UNIPROT:P01277; UNIPARC:UPI000012B82E; GB:S78477; NID:g999386; PIDN:
C;Superfamily: glucagon
C;Keywords: duplication

  Query Match             48.5%;  Score 225.5;  DB 2;  Length 206;
  Best Local Similarity   37.1%;  Pred. No. 6.1e-16;
  Matches   53;  Conservative   15;  Mismatches   20;  Indels   55;  Gaps    4;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRG---------SLD--------KRHGEGTFTSDV 39
              :|::|    ||  |||      |: |  :           ||        ||| :||||||
Db          5 SIYFIAGLLLMIVQGSWQNPLQDTEEKSRSFKASQSEPLDESRQLNEVKRHSQGTFTSDY 64

Qy         40 SSYLEGQAAKEFIAWLVK--------------------------------GRHGEGTFT 66
              | ||: : |::|: ||:                                ||  |||:|
Db         65 SKYLDSRRAQDFVQWLMSTKRNGQQGQEDKENDKFPDQLSSNAISKRHSEFERHAEGTYT 124

Qy         67 SDVSSYLEGQAAKEFIAWLVKGR 89
              ||::|||||||||||||||| ||
Db        125 SDITSYLEGQAAKEFIAWLVNGR 147


RESULT 11
I51058
glucagon I precursor - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51058; I51299; I51056; I51037; I51036; I51300
R;Irwin, D.M.; Wong, J.
Mol. Endocrinol. 9, 267-277, 1995
A;Title: Trout and chicken proglucagon: alternative splicing generates mRNA transcripts
A;Reference number: A55895; MUID:95295739; PMID:7776976
A;Accession: I51058
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-178 <IRW>
A;Cross-references: UNIPROT:Q91971; UNIPARC:UPI00000FE622; EMBL:U19917; NID:g736364; PID
A;Accession: I51299
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 52-53,'X',55-123 <IR2>
A;Cross-references: UNIPARC:UPI0000176628; GB:S78473; NID:g999382; PIDN:AAB34504.1; PID:
A;Accession: I51056
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 58-123 <IR3>
A;Cross-references: UNIPARC:UPI000017142D; EMBL:U19913; NID:g736360; PIDN:AAC59667.1; PI
A;Accession: I51037
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 'M',114-144 <IR4>
A;Cross-references: UNIPARC:UPI0000176629; EMBL:U19919; NID:g736374; PIDN:AAC60213.1; PI
A;Accession: I51036
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 113-123 <IR5>
A;Cross-references: UNIPARC:UPI0000171434; EMBL:U19918; NID:g736373; PIDN:AAC60212.1; PI
C;Genetics:
A;Introns: 123/2
C;Superfamily: glucagon
C;Keywords: duplication

  Query Match             44.8%;  Score 208.5;  DB 2;  Length 178;
  Best Local Similarity   42.6%;  Pred. No. 2.9e-14;
  Matches   40;  Conservative   19;  Mismatches   18;  Indels   17;  Gaps    1;

Qy         13 FVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKG---------- 58
              ||| | :: | |:  ||| :||:|||||:||: ||||:|::||  |
Db         73 FVQWLMNSKRSGAPSKRHADGTYTSDVSTYLQDQAAKDFVSWLKSGRARRESAEESRNGP 132

Qy         59 ---RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
                 || :|:|||||:  |:  ||||:: |::  :
Db        133 MSRRHVDGSFTSDVNKVLDSLAAKEYLLWVMTSK 166


RESULT 12
GCFGB
glucagon precursor - bullfrog (fragments)
N;Alternate names: oxyntomodulin
N;Contains: glucagon; glucagon-36 (oxyntomodulin); glucagon-like peptide 1; glucagon-li
C;Species: Rana catesbeiana (bullfrog)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 20-Mar-1998
C;Accession: B28091; C28091; D28091
R;Pollock, H.G.; Hamilton, J.W.; Rouse, J.B.; Ebner, K.E.; Rawitch, A.B.
J. Biol. Chem. 263, 9746-9751, 1988
A;Title: Isolation of peptide hormones from the pancreas of the bullfrog (Rana catesbei
A;Reference number: A92730; MUID:88257102; PMID:3260236
A;Accession: B28091
A;Molecule type: protein
A;Residues: 1-36 <PO2>
A;Cross-references: UNIPARC:UPI0000173502
A;Accession: C28091
A;Molecule type: protein
A;Residues: 37-68 <POL>
A;Cross-references: UNIPARC:UPI0000173502
A;Accession: D28091
A;Molecule type: protein
A;Residues: 69-101 <PO3>
A;Cross-references: UNIPARC:UPI0000173502
C;Superfamily: glucagon
C;Keywords: carbohydrate metabolism; duplication; hormone; pancreas
F;1-36/Product: glucagon-36 (oxyntomodulin) #status experimental <G36>
F;1-29/Product: glucagon #status predicted <GCN>
F;37-67/Product: glucagon-like peptide 1 #status experimental <GL1>
F;69-101/Product: glucagon-like peptide 2 #status experimental <GL2>

  Query Match             44.3%;  Score 206;  DB 1;  Length 101;
  Best Local Similarity   52.6%;  Pred. No. 2.9e-14;
  Matches   40;  Conservative   17;  Mismatches   15;  Indels    4;  Gaps    2;

Qy         13 FVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR--HGEGTFTSDVS 70
              ||| | :: | | :   | :||||||:||||| :|||||: ||:|||  | :|:|||| :
Db         22 FVQWLMNSKRSGGIS--HADGTFTSDMSSYLEEKAAKEFVDWLIKGRPKHADGSFTSDFN 79

Qy         71 SYLEGQAAKEFIAWLV 86
                |: :||:||: |::
Db         80 KALDIKAAQEFLDWII 95


RESULT 13
GCAF2
glucagon 2 precursor - American goosefish
```

```
N;Contains: glucagon; glucagon-like peptide 1
C;Species: Lophius americanus (American goosefish)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: A05150
R;Lund, P.K.; Goodman, R.H.; Montminy, M.R.; Dee, P.C.; Habener, J.F.
J. Biol. Chem. 258, 3280-3284, 1983
A;Title: Anglerfish islet pre-proglucagon II. Nucleotide and corresponding amino acid se
A;Reference number: A05150; MUID:83135785; PMID:6338015
A;Accession: A05150
A;Molecule type: mRNA
A;Residues: 1-122 <LUN>
A;Cross-references: UNIPROT:P04092; UNIPARC:UPI000012B81E; GB:J00933; NID:g64021; PIDN:C
C;Superfamily: glucagon
C;Keywords: carbohydrate metabolism; duplication; hormone; pancreas
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-122/Product: proglucagon 2 #status predicted <PGC2>
F;52-80/Product: glucagon #status predicted <GCN>
F;89-119/Product: glucagon-like peptide 1 #status predicted <GL1>

  Query Match             43.1%;  Score 200.5;  DB 1;  Length 122;
  Best Local Similarity   53.6%;  Pred. No. 1.3e-13;
  Matches   37;  Conservative   12;  Mismatches   13;  Indels    7;  Gaps    1;

Qy         28 KRHGEGTFTSDVSSYLEGQAAKEFIAWLVKG-------RHGEGTFTSDVSSYLEGQAAKE 80
              ||| |||| ::| | ||| : |::|: ||          || :||:||||||||: ||||:
Db         50 KRHSEGTFSNDYSKYLETRRAQDFVQWLKNSKRNGLFRRHADGTYTSDVSSYLQDQAAKD 109

Qy         81 FIAWLVKGR 89
              |::||  ||
Db        110 FVSWLKAGR 118


RESULT 14
I51057
glucagon II precursor - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51057; I51039; I51038
R;Irwin, D.M.; Wong, J.
Mol. Endocrinol. 9, 267-277, 1995
A;Title: Trout and chicken proglucagon: alternative splicing generates mRNA transcripts
A;Reference number: A55895; MUID:95295739; PMID:7776976
A;Accession: I51057
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-178 <IRW>
A;Cross-references: UNIPROT:Q91189; UNIPARC:UPI00000FE623; EMBL:U19914; NID:g736362; PID
A;Accession: I51039
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 113-144 <IR2>
A;Cross-references: UNIPARC:UPI0000171437; EMBL:U19916; NID:g736369; PIDN:AAC60210.1; PI
A;Accession: I51038
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 113-123 <IR3>
A;Cross-references: UNIPARC:UPI0000171436; EMBL:U19915; NID:g736368; PIDN:AAC60209.1; PI
C;Genetics:
A;Introns: 123/2
C;Superfamily: glucagon
C;Keywords: duplication

  Query Match             43.1%;  Score 200.5;  DB 2;  Length 178;
  Best Local Similarity   40.4%;  Pred. No. 1.9e-13;
  Matches   38;  Conservative   20;  Mismatches   19;  Indels   17;  Gaps    1;

Qy         13 FVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKG---------- 58
              |:  | :: | |:  ||| :||:|||||:||: ||||:|::||  |
Db         73 FLHWLMNSKRSGAPSKRHADGTYTSDVSTYLQDQAAKDFVSWLKSGPARRESAEESMNGP 132

Qy         59 ---RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
                 || :|:|||||:  |:  ||||:: |::  :
Db        133 MSRRHVDGSFTSDVNKVLDSLAAKEYLLWVMTSK 166


RESULT 15
GCIDC
glucagon precursor - channel catfish (fragments)
C;Species: Ictalurus punctatus (channel catfish)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: A05166; A05167
R;Andrews, P.C.; Ronner, P.
J. Biol. Chem. 260, 3910-3914, 1985
A;Title: Isolation and structures of glucagon and glucagon-like peptide from catfish pan
A;Reference number: A92514; MUID:85157536; PMID:3838546
A;Accession: A05166
A;Molecule type: protein
A;Residues: 1-29 <AND1>
A;Cross-references: UNIPROT:P04093; UNIPARC:UPI0000173508
A;Accession: A05167
A;Molecule type: protein
A;Residues: 30-63 <AND2>
A;Cross-references: UNIPARC:UPI0000173509
C;Superfamily: glucagon
C;Keywords: carbohydrate metabolism; duplication; hormone; pancreas
F;1-29/Product: glucagon #status experimental <GCN>
F;30-63/Product: glucagon-like peptide 1 #status experimental <GL1>

  Query Match             40.5%;  Score 188.5;  DB 1;  Length 63;
  Best Local Similarity   56.7%;  Pred. No. 1.1e-12;
  Matches   34;  Conservative   12;  Mismatches   13;  Indels    1;  Gaps    1;

Qy         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              | ||||::| | ||| : |::|: ||:   | :||:||||||||: ||||:|| ||  |:
Db          1 HSEGTFSNDYSKYLETRRAQDFVQWLMNS-HADGTYTSDVSSYLQDQAAKDFITWLKSGQ 59


Search completed: April 19, 2006, 12:10:00
Job time : 4.04545 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 11:57:02 ; Search time 23.5142 Seconds
(without alignments)
2670.387 Million cell updates/sec

Title: US-10-775-180-449
Perfect score: 465
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....SSYLEGQAAKEFIAWLVKGR 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	251	54.0	220	2	Q8UWL9_9NEOB	Q8uwl9 hoplobatrac
2	249.5	53.7	180	1	GLUC_MESAU	P01273 m glucagon
3	247.5	53.2	176	1	GLUC_SHEEP	Q8mj25 o glucagon
4	247.5	53.2	180	1	GLUC_BOVIN	P01272 b glucagon
5	244.5	52.6	180	1	GLUC_PIG	P01274 s glucagon
6	243.5	52.4	180	1	GLUC_CANFA	P29794 c glucagon
7	243.5	52.4	180	1	GLUC_MOUSE	P55095 m glucagon
8	243.5	52.4	180	2	Q53TP6_HUMAN	Q53tp6 homo sapien
9	241.5	51.9	180	1	GLUC_HUMAN	P01275 h glucagon
10	241.5	51.9	180	1	GLUC_RAT	P06883 r glucagon
11	240.5	51.7	180	1	GLUC_CAVPO	P05110 c glucagon
12	238	51.2	149	2	Q6RYB2_BUFMA	Q6ryb2 bufo marinu
13	238	51.2	266	1	GLUC1_XENLA	O42143 xenopus lae
14	233.5	50.2	266	2	Q6DIZ4_XENTR	Q6diz4 xenopus tro
15	230.5	49.6	180	1	GLUC_OCTDE	P22890 o glucagon
16	229.5	49.4	219	1	GLUC2_XENLA	O42144 xenopus lae
17	229.5	49.4	219	2	Q5D082_XENLA	Q5d082 xenopus lae
18	226	48.6	145	2	Q6RYB5_NEOFS	Q6ryb5 neoceratodu
19	225.5	48.5	206	1	GLUC_CHICK	P68259 g glucagon
20	222	47.7	153	2	Q6RYB6_PRODO	Q6ryb6 protopterus
21	215	46.2	103	1	GLUC_RANCA	P15438 rana catesb
22	215	46.2	124	2	Q6RYB1_9SAUR	Q6ryb1 agkistrodon
23	214.5	46.1	204	1	GLUC_HELSU	O12956 h glucagon
24	208.5	44.8	178	1	GLUC1_ONCMY	Q91971 oncorhynchu
25	205.5	44.2	120	2	Q6RYB7_ICTPU	Q6ryb7 ictalurus p
26	203	43.7	160	1	GLUC1_PETMA	Q9pur1 petromyzon
27	201.5	43.3	173	2	Q6RYB9_ICTPU	Q6ryb9 ictalurus p
28	200.5	43.1	122	1	GLUC2_LOPAM	P04092 lophius ame
29	200.5	43.1	178	1	GLUC2_ONCMY	Q91189 oncorhynchu
30	200	43.0	121	2	Q5PR39_BRARE	Q5pr39 brachydanio
31	200	43.0	123	2	Q6RYA9_9PERC	Q6rya9 sebastes ca
32	200	43.0	860	2	Q4RQJ4_TETNG	Q4rqj4 tetraodon n
33	198	42.6	121	1	GLUC_CARAU	P79695 carassius a
34	197	42.4	122	2	Q6RYB8_ICTPU	Q6ryb8 ictalurus p
35	193	41.5	124	2	Q4S308_TETNG	Q4s308 tetraodon n
36	189.5	40.8	176	2	Q6RYB0_9PERC	Q6ryb0 sebastes ca
37	189.5	40.8	176	2	Q6RYC2_9PERC	Q6ryc2 sebastes ca
38	189	40.6	170	2	Q6RYB4_SQUAC	Q6ryb4 squalus aca
39	186	40.0	80	2	Q6IUP8_PHOSU	Q6iup8 phodopus su
40	186	40.0	121	2	Q9DDE6_BRARE	Q9dde6 brachydanio
41	185.5	39.9	71	1	GLUC_PIAME	P81880 piaractus m
42	183.5	39.5	71	1	GLUC_ICTPU	P04093 ictalurus p
43	183	39.4	124	1	GLUC1_LOPAM	P01278 lophius ame
44	179.5	38.6	121	2	Q6RYC1_9PERC	Q6ryc1 sebastes ca
45	179	38.5	78	1	GLUC_LEPSP	P09566 lepisosteus

ALIGNMENTS

RESULT 1
Q8UWL9_9NEOB
ID Q8UWL9_9NEOB PRELIMINARY; PRT; 220 AA.
AC Q8UWL9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Proglucagon.
OS Hoplobatrachus rugulosus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae;
OC Hoplobatrachus.
OX NCBI_TaxID=110072;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yeung C.-M., Chow B.K.C.;
RT "Identification of a proglucagon cDNA from Rana tigrina rugulosa that
RT encodes two GLP-1s.";
RL Gen. Comp. Endocrinol. 124:0-0(2001).
DR EMBL; AF324209; AAL35758.1; -; mRNA.
DR HSSP; P01274; 1GCN.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000532; Glucagon.
DR Pfam; PF00123; Hormone_2; 4.
DR PRINTS; PR00275; GLUCAGON.
DR SMART; SM00070; GLUCA; 4.
DR PROSITE; PS00260; GLUCAGON; 4.
SQ SEQUENCE 220 AA; 25615 MW; C72D926E7F89E381 CRC64;

Query Match 54.0%; Score 251; DB 2; Length 220;
Best Local Similarity 54.9%; Pred. No. 2.9e-18;
Matches 50; Conservative 14; Mismatches 13; Indels 14; Gaps 2;

```
Qy         13 FVQGLEHTHRRGSLDK------RHGEGTFTSDVSSYLEGQAAKEFIAWLVKG-------- 58
              ||| | :: | | : |      || |||:|:||: :|| :|||||| ||:||
Db         74 FVQWLMNSKRSGGISKRNVQFERHAEGTYTNDVTQFLEEKAAKEFIDWLIKGKPKKQRLS 133

Qy         59 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              || |||||||::|||| :|||||: ||:|||
Db        134 RHAEGTFTSDMTSYLEEKAAKEFVDWLIKGR 164
```

RESULT 2
GLUC_MESAU
ID GLUC_MESAU STANDARD; PRT; 180 AA.
AC P01273;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)].

```
GN   Name=GCG;
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Cricetidae; Cricetinae; Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=83167563; PubMed=6835407;
RA   Bell G.I., Santerre R.F., Mullenbach G.T.;
RT   "Hamster preproglucagon contains the sequence of glucagon and two
RT   related peptides.";
RL   Nature 302:716-718(1983).
RN   [2]
RP   SEQUENCE REVISION TO 12-15.
RA   Bell G.I.;
RL   Submitted (JUN-1985) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   REVIEW.
RX   MEDLINE=22442611; PubMed=12554744; DOI=10.1210/me.2002-0306;
RA   Drucker D.J.;
RT   "Glucagon-like peptides: regulators of cell proliferation,
RT   differentiation, and apoptosis.";
RL   Mol. Endocrinol. 17:161-171(2003).
RN   [4]
RP   REVIEW.
RX   MEDLINE=22513095; PubMed=12626323; DOI=10.1152/ajpendo.00492.2002;
RA   Jiang G., Zhang B.B.;
RT   "Glucagon and regulation of glucose metabolism.";
RL   Am. J. Physiol. 284:E671-E678(2003).
RN   [5]
RP   REVIEW.
RX   PubMed=10322410;
RA   Drucker D.J.;
RT   "Glucagon-like peptide 2.";
RL   Trends Endocrinol. Metab. 10:153-156(1999).
RN   [6]
RP   REVIEW.
RX   MEDLINE=20073561; PubMed=10605628; DOI=10.1210/er.20.6.876;
RA   Kieffer T.J., Habener J.F.;
RT   "The glucagon-like peptides.";
RL   Endocr. Rev. 20:876-913(1999).
CC   -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC       homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC       and decreasing glycolysis. A counterregulatory hormone of insulin,
CC       raises plasma glucose levels in response to insulin-induced
CC       hypoglycemia (By similarity).
CC   -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC       insulin release. Play important roles on gastric motility and the
CC       suppression of plasma glucagon levels. May be involved in the
CC       suppression of satiety and stimulation of glucose disposal in
CC       peripheral tissues, independent of the actions of insulin. Have
CC       growth-promoting activities on intestinal epithelium. May also
CC       regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC       TSH, CRH, oxytocin, and vasopressin secretion. Increases islet
CC       mass through stimulation of islet neogenesis and pancreatic beta
CC       cell proliferaton (By similarity).
CC   -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC       villus height in the small intestine, concomitant with increased
CC       crypt cell proliferation and decreased enterocyte apoptosis. The
CC       gastrointestinal tract, from the stomach to the colon is the
CC       principal target for GLP-2 action. Plays a key role in nutrient
CC       homeostasis, enhancing nutrient assimilation through enhanced
CC       gastrointestinal function, as well as increasing nutrient
CC       disposal. Stimulates intestinal glucose transport and decreases
CC       mucosal permeability (By similarity).
CC   -!- FUNCTION: Oxyntomodulin significantly reduces food intake (By
CC       similarity).
CC   -!- FUNCTION: Glicentin may modulate gastric acid secretion and
CC       gastro-pyloro-duodenal activity (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- INDUCTION: Glucagon release is stimulated by hypoglycemia and
CC       inhibited by hyperglycemia, insulin, and somatostatin. GLP-1 and
CC       GLP-2 are induced in response to nutrient ingestion (By
CC       similarity).
CC   -!- PTM: Proglucagon is posttranslationally processed in a tissue-
CC       specific manner in pancreatic A cells and intestinal L cells. In
CC       pancreatic A cells, the major bioactive hormone is glucagon
CC       cleaved by PCSK2/PC2. In the intestinal L cells PCSK1/PC1
CC       liberates GLP-1, GLP-2, glicentin and oxyntomodulin. GLP-1 is
CC       further N-terminally truncated by posttranslational processing in
CC       the intestinal L cells resulting in GLP-1(7-37) GLP-1-(7-36)amide.
CC       The C-terminal amidation is neither important for the metabolism
CC       of GLP-1 nor for its effects on the endocrine pancreas (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; J00059; AAA37074.1; -; mRNA.
DR   HSSP; P01275; 1D0R.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Amidation; Cleavage on pair of basic residues; Glucagon family;
KW   Hormone; Signal.
FT   SIGNAL       1     20
FT   PEPTIDE      21    89       Glicentin (By similarity).
FT   PEPTIDE      21    50       Glicentin-related polypeptide (By
FT                               similarity).
FT   PEPTIDE      53    89       Oxyntomodulin (By similarity).
FT   PEPTIDE      53    81       Glucagon (By similarity).
FT   PROPEP       84    89       By similarity.
FT   PEPTIDE      92    128      Glucagon-like peptide 1 (By similarity).
FT   PEPTIDE      98    128      Glucagon-like peptide 1(7-37) (By
FT                               similarity).
FT   PEPTIDE      98    127      Glucagon-like peptide 1(7-36) (By
FT                               similarity).
FT   PROPEP       131   145      By similarity.
FT   PEPTIDE      146   178      Glucagon-like peptide 2 (By similarity).
FT   SITE         52    53       Cleavage (by PCSK2) (By similarity).
FT   SITE         83    84       Cleavage (by PCSK1 and PCSK2) (By
FT                               similarity).
FT   SITE         91    92       Cleavage (by PCSK1) (By similarity).
FT   SITE         97    98       Cleavage (by PCSK1) (By similarity).
FT   SITE         130   131      Cleavage (by PCSK1) (By similarity).
FT   SITE         145   146      Cleavage (by PCSK1) (By similarity).
FT   MOD_RES      127   127      Arginine amide (G-128 provides amide
FT                               group) (By similarity).
SQ   SEQUENCE   180 AA;  20954 MW;  02791B49D7AADD4B CRC64;

  Query Match             53.7%;  Score 249.5;  DB 1;  Length 180;
  Best Local Similarity   44.0%;  Pred. No. 3.3e-18;
  Matches   55;  Conservative   13;  Mismatches   20;  Indels   37;  Gaps    3;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRG----------------SLDKRHGEGTFTSDV 39
              ||: :  | :  |||     |: |  :                : |||| :||||||
Db          3 NIYIVAGFFVVLVQGSWQHSLQDTEEKSRSFPASQTDPLEDPDQINEDKRHSQGTFTSDY 62

Qy         40 SSYLEGQAAKEFIAWLVK---------------GRHGEGTFTSDVSSYLEGQAAKEFIAW 84
               | ||: : |::|: ||:                || ||||||||||||||||||||||
Db         63 SKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAW 122

Qy         85 LVKGR 89
              |||||
Db        123 LVKGR 127


RESULT 3
GLUC_SHEEP
ID   GLUC_SHEEP     STANDARD;      PRT;   176 AA.
```

```
AC   Q8MJ25;
DT   25-OCT-2004 (Rel. 45, Created)
DT   25-OCT-2004 (Rel. 45, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE   (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE   (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE   peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)]
DE   (Fragment).
GN   Name=GCG;
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC   Pecora; Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Pancreas;
RA   Limesand S.W., Hay W.W. Jr.;
RT   "Characterization of the endocrine pancreas in an ovine placental
RT   insufficiency IUGR fetus.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC       homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC       and decreasing glycolysis. A counterregulatory hormone of insulin,
CC       raises plasma glucose levels in response to insulin-induced
CC       hypoglycemia (By similarity).
CC   -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC       insulin release. Play important roles on gastric motility and the
CC       suppression of plasma glucagon levels. May be involved in the
CC       suppression of satiety and stimulation of glucose disposal in
CC       peripheral tissues, independent of the actions of insulin. Have
CC       growth-promoting activities on intestinal epithelium. May also
CC       regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC       TSH, CRH, oxytocin, and vasopressin (By similarity).
CC   -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC       villus height in the small intestine, concomitant with increased
CC       crypt cell proliferation and decreased enterocyte apoptosis. The
CC       gastrointestinal tract, from the stomach to the colon is the
CC       principal target for GLP-2 action. Plays a key role in nutrient
CC       homeostasis, enhancing nutrient assimilation through enhanced
CC       gastrointestinal function, as well as increasing nutrient
CC       disposal. Stimulates intestinal glucose transport and decreases
CC       mucosal permeability (By similarity).
CC   -!- FUNCTION: Oxyntomodulin significantly reduces food intake (By
CC       similarity).
CC   -!- FUNCTION: Glicentin may modulate gastric acid secretion and
CC       gastro-pyloro-duodenal activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Glucagon is secreted in the A cells of the
CC       islets of Langerhans. GLP-1, GLP-2, oxyntomodulin and glicentin
CC       are secreted from enteroendocrine cells throughout the
CC       gastrointestinal tract. GLP1 and GLP2 are also secreted in
CC       selected neurons in the brain.
CC   -!- INDUCTION: Glucagon release is stimulated by hypoglycemia and
CC       inhibited by hyperglycemia, insulin, and somatostatin. GLP-1 and
CC       GLP-2 are induced in response to nutrient ingestion (By
CC       similarity).
CC   -!- PTM: Proglucagon is posttranslationally processed in a tissue-
CC       specific manner in pancreatic A cells and intestinal L cells. In
CC       pancreatic A cells, the major bioactive hormone is glucagon
CC       cleaved by PCSK2/PC2. In the intestinal L cells PCSK1/PC1
CC       liberates GLP-1, GLP-2, glicentin and oxyntomodulin. GLP-1 is
CC       further N-terminally truncated by posttranslational processing in
CC       the intestinal L cells resulting in GLP-1(7-37) GLP-1-(7-36)amide.
CC       The C-terminal amidation is neither important for the metabolism
CC       of GLP-1 nor for its effects on the endocrine pancreas (By
CC       similarity).
CC   -!- MISCELLANEOUS: GLP-2 does not have cleavage on a pair of basic
CC       residues at C-terminus as in other mammmals.
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF529185; AAM94409.1; -; mRNA.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Amidation; Cleavage on pair of basic residues; Glucagon family;
KW   Hormone; Signal.
FT   SIGNAL       1     20
FT   PEPTIDE     21     89       Glicentin (By similarity).
FT   PEPTIDE     21     50       Glicentin-related polypeptide (By
FT                               similarity).
FT   PEPTIDE     53     89       Oxyntomodulin (By similarity).
FT   PEPTIDE     53     81       Glucagon.
FT   PROPEP      84     89       By similarity.
FT   PEPTIDE     92    128       Glucagon-like peptide 1 (By similarity).
FT   PEPTIDE     98    128       Glucagon-like peptide 1(7-37) (By
FT                               similarity).
FT   PEPTIDE     98    127       Glucagon-like peptide 1(7-36) (By
FT                               similarity).
FT   PROPEP     131    145       By similarity.
FT   PEPTIDE    146   >176       Glucagon-like peptide 2 (By similarity).
FT   SITE        52     53       Cleavage (by PCSK2) (By similarity).
FT   SITE        83     84       Cleavage (by PCSK1 and PCSK2) (By
FT                               similarity).
FT   SITE        91     92       Cleavage (by PCSK1) (By similarity).
FT   SITE        97     98       Cleavage (by PCSK1) (By similarity).
FT   SITE       130    131       Cleavage (by PCSK1) (By similarity).
FT   SITE       145    146       Cleavage (by PCSK1) (By similarity).
FT   MOD_RES    127    127       Arginine amide (G-128 provides amide
FT                               group) (By similarity).
FT   NON_TER    176    176
SQ   SEQUENCE   176 AA;  20336 MW;  13174039BD6CE2B3 CRC64;

  Query Match             53.2%;  Score 247.5;  DB 1;  Length 176;
  Best Local Similarity   42.4%;  Pred. No. 5.3e-18;
  Matches   53;  Conservative   16;  Mismatches   19;  Indels   37;  Gaps    3;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRGSL----------------DKRHGEGTFTSDV 39
              :::::   |:   ||      |::|  : |                  |||| :||||||
Db          3 SLYFVAGLLVMLAQGSWQHSLQNTEEKSSSFPAPQTDPLGDPDQISEDKRHSQGTFTSDY 62

Qy         40 SSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSSYLEGQAAKEFIAW 84
              | ||: : |::|: ||:                || ||||||||||||||||||||
Db         63 SKYLDSRRAQDFVQWLMNTKRNKNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAW 122

Qy         85 LVKGR 89
              |||||
Db        123 LVKGR 127


RESULT 4
GLUC_BOVIN
ID   GLUC_BOVIN     STANDARD;      PRT;   180 AA.
AC   P01272;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE   (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE   (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE   peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)].
GN   Name=GCG;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC   Pecora; Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
```

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=83299996; PubMed=6577439;
RA   Lopez L.C., Frazier M.L., Su C.-J., Kumar A., Saunders G.F.;
RT   "Mammalian pancreatic preproglucagon contains three glucagon-related
RT   peptides.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:5485-5489(1983).
RN   [2]
RP   PROTEIN SEQUENCE OF 53-81.
RX   MEDLINE=71166445; PubMed=5102927;
RA   Bromer W.W., Boucher M.E., Koffenberger J.E. Jr.;
RT   "Amino acid sequence of bovine glucagon.";
RL   J. Biol. Chem. 246:2822-2827(1971).
RN   [3]
RP   REVIEW.
RX   MEDLINE=22442611; PubMed=12554744; DOI=10.1210/me.2002-0306;
RA   Drucker D.J.;
RT   "Glucagon-like peptides: regulators of cell proliferation,
RT   differentiation, and apoptosis.";
RL   Mol. Endocrinol. 17:161-171(2003).
RN   [4]
RP   REVIEW.
RX   MEDLINE=22513095; PubMed=12626323; DOI=10.1152/ajpendo.00492.2002;
RA   Jiang G., Zhang B.B.;
RT   "Glucagon and regulation of glucose metabolism.";
RL   Am. J. Physiol. 284:E671-E678(2003).
RN   [5]
RP   REVIEW.
RX   PubMed=10322410;
RA   Drucker D.J.;
RT   "Glucagon-like peptide 2.";
RL   Trends Endocrinol. Metab. 10:153-156(1999).
RN   [6]
RP   REVIEW.
RX   MEDLINE=20073561; PubMed=10605628; DOI=10.1210/er.20.6.876;
RA   Kieffer T.J., Habener J.F.;
RT   "The glucagon-like peptides.";
RL   Endocr. Rev. 20:876-913(1999).
RN   [7]
RP   STRUCTURE BY NMR OF 53-81.
RX   MEDLINE=71166445; PubMed=6631957;
RA   Braun W., Wider G., Lee K.H., Wuethrich K.;
RT   "Conformation of glucagon in a lipid-water interphase by 1H nuclear
RT   magnetic resonance.";
RL   J. Mol. Biol. 169:921-948(1983).
CC   -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC       homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC       and decreasing glycolysis. A counterregulatory hormone of insulin,
CC       raises plasma glucose levels in response to insulin-induced
CC       hypoglycemia (By similarity).
CC   -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC       insulin release. Play important roles on gastric motility and the
CC       suppression of plasma glucagon levels. May be involved in the
CC       suppression of satiety and stimulation of glucose disposal in
CC       peripheral tissues, independent of the actions of insulin. Have
CC       growth-promoting activities on intestinal epithelium. May also
CC       regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC       TSH, CRH, oxytocin, and vasopressin secretion. Increases islet
CC       mass through stimulation of islet neogenesis and pancreatic beta
CC       cell proliferaton (By similarity).
CC   -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC       villus height in the small intestine, concomitant with increased
CC       crypt cell proliferation and decreased enterocyte apoptosis. The
CC       gastrointestinal tract, from the stomach to the colon is the
CC       principal target for GLP-2 action. Plays a key role in nutrient
CC       homeostasis, enhancing nutrient assimilation through enhanced
CC       gastrointestinal function, as well as increasing nutrient
CC       disposal. Stimulates intestinal glucose transport and decreases
CC       mucosal permeability (By similarity).
CC   -!- FUNCTION: Oxyntomodulin significantly reduces food intake (By
CC       similarity).
CC   -!- FUNCTION: Glicentin may modulate gastric acid secretion and
CC       gastro-pyloro-duodenal activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Glucagon is secreted in the A cells of the
CC       islets of Langerhans. GLP-1, GLP-2, oxyntomodulin and glicentin
CC       are secreted from enteroendocrine cells throughout the
CC       gastrointestinal tract.
CC   -!- INDUCTION: Glucagon release is stimulated by hypoglycemia and
CC       inhibited by hyperglycemia, insulin, and somatostatin. GLP-1 and
CC       GLP-2 are induced in response to nutrient ingestion (By
CC       similarity).
CC   -!- PTM: Proglucagon is posttranslationally processed in a tissue-
CC       specific manner in pancreatic A cells and intestinal L cells. In
CC       pancreatic A cells, the major bioactive hormone is glucagon
CC       cleaved by PCSK2/PC2. In the intestinal L cells PCSK1/PC1
CC       liberates GLP-1, GLP-2, glicentin and oxyntomodulin. GLP-1 is
CC       further N-terminally truncated by posttranslational processing in
CC       the intestinal L cells resulting in GLP-1(7-37) GLP-1-(7-36)amide.
CC       The C-terminal amidation is neither important for the metabolism
CC       of GLP-1 nor for its effects on the endocrine pancreas (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; K00107; AAA30538.1; -; mRNA.
DR   PDB; 1KX6; NMR; A=53-81.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   3D-structure; Amidation; Cleavage on pair of basic residues;
KW   Direct protein sequencing; Glucagon family; Hormone; Signal.
FT   SIGNAL       1     20
FT   PEPTIDE     21     89       Glicentin (By similarity).
FT   PEPTIDE     21     50       Glicentin-related polypeptide (By
FT                               similarity).
FT   PEPTIDE     53     89       Oxyntomodulin (By similarity).
FT   PEPTIDE     53     81       Glucagon.
FT   PROPEP      84     89       By similarity.
FT   PEPTIDE     92    128       Glucagon-like peptide 1 (By similarity).
FT   PEPTIDE     98    128       Glucagon-like peptide 1(7-37) (By
FT                               similarity).
FT   PEPTIDE     98    127       Glucagon-like peptide 1(7-36) (By
FT                               similarity).
FT   PROPEP     131    145       By similarity.
FT   PEPTIDE    146    178       Glucagon-like peptide 2 (By similarity).
FT   SITE        52     53       Cleavage (by PCSK2) (By similarity).
FT   SITE        83     84       Cleavage (by PCSK1 and PCSK2) (By
FT                               similarity).
FT   SITE        91     92       Cleavage (by PCSK1) (By similarity).
FT   SITE        97     98       Cleavage (by PCSK1) (By similarity).
FT   SITE       130    131       Cleavage (by PCSK1) (By similarity).
FT   SITE       145    146       Cleavage (by PCSK1) (By similarity).
FT   MOD_RES    127    127       Arginine amide (G-128 provides amide
FT                               group) (By similarity).
FT   TURN        60     64
FT   TURN        74     74
FT   HELIX       75     78
SQ   SEQUENCE   180 AA;  20944 MW;  8D9B4FF05B9F15FF CRC64;

  Query Match             53.2%;  Score 247.5;  DB 1;  Length 180;
  Best Local Similarity   42.4%;  Pred. No. 5.4e-18;
  Matches   53;  Conservative   16;  Mismatches   19;  Indels   37;  Gaps    3;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRGSL----------------DKRHGEGTFTSDV 39
              :::::    :  |||     |::|  : |                 |||| :||||||
Db          3 SLYFVAGLFVMLVQGSWQRSLQNTEEKSSSFPAPQTDPLGDPDQINEDKRHSQGTFTSDY 62

Qy         40 SSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSSYLEGQAAKEFIAW 84
               | ||: : |::|: ||:                || ||||||||||||||||||||||
```

```
Db         63 SKYLDSRRAQDFVQWLMNTKRNKNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAW 122

Qy         85 LVKGR 89
              |||||
Db        123 LVKGR 127


RESULT 5
GLUC_PIG
ID   GLUC_PIG       STANDARD;      PRT;   180 AA.
AC   P01274; Q864V8;
DT   21-JUL-1986 (Rel. 01, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE   (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE   (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE   peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)].
GN   Name=GCG;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC   Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Pancreas, and Small intestine;
RA   Siggers R.H., Goldade B.G., Laarveld B., Van Kessel A.G.;
RT   "Cloning of porcine proglucagon.";
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   PROTEIN SEQUENCE OF 21-89.
RX   MEDLINE=81248172; PubMed=6894800; DOI=10.1016/0167-0115(81)90007-0;
RA   Thim L., Moody A.J.;
RT   "The primary structure of porcine glicentin (proglucagon).";
RL   Regul. Pept. 2:139-150(1981).
RN   [3]
RP   PROTEIN SEQUENCE OF 21-89.
RX   MEDLINE=82221776; PubMed=7045833; DOI=10.1016/0196-9781(81)90007-3;
RA   Thim L., Moody A.J.;
RT   "The amino acid sequence of porcine glicentin.";
RL   Peptides 2 Suppl. 2:37-39(1981).
RN   [4]
RP   PROTEIN SEQUENCE OF 53-81.
RA   Bromer W.W., Sinn L.G., Behrens O.K.;
RT   "The amino acid sequence of glucagon. V. Location of amide groups,
RT   acid degradation studies and summary of sequential evidence.";
RL   J. Am. Chem. Soc. 79:2807-2810(1957).
RN   [5]
RP   PROTEIN SEQUENCE OF 98-127.
RX   MEDLINE=89327238; PubMed=2753890;
RA   Orskov C., Bersani M., Johnsen A.H., Hoejrup P., Holst J.J.;
RT   "Complete sequences of glucagon-like peptide-1 from human and pig
RT   small intestine.";
RL   J. Biol. Chem. 264:12826-12829(1989).
RN   [6]
RP   PROTEIN SEQUENCE OF 131-178.
RX   MEDLINE=88243712; PubMed=3379036;
RA   Buhl T., Thim L., Kofod H., Orskov C., Harling H., Holst J.J.;
RT   "Naturally occurring products of proglucagon 111-160 in the porcine
RT   and human small intestine.";
RL   J. Biol. Chem. 263:8621-8624(1988).
RN   [7]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=87004290; PubMed=3530719;
RA   Orskov C., Holst J.J., Knuhtsen S., Baldissera F.G., Poulsen S.S.,
RA   Nielsen O.V.;
RT   "Glucagon-like peptides GLP-1 and GLP-2, predicted products of the
RT   glucagon gene, are secreted separately from pig small intestine but
RT   not pancreas.";
RL   Endocrinology 119:1467-1475(1986).
RN   [8]
RP   REVIEW.
RX   MEDLINE=22442611; PubMed=12554744; DOI=10.1210/me.2002-0306;
RA   Drucker D.J.;
RT   "Glucagon-like peptides: regulators of cell proliferation,
RT   differentiation, and apoptosis.";
RL   Mol. Endocrinol. 17:161-171(2003).
RN   [9]
RP   REVIEW.
RX   MEDLINE=22513095; PubMed=12626323; DOI=10.1152/ajpendo.00492.2002;
RA   Jiang G., Zhang B.B.;
RT   "Glucagon and regulation of glucose metabolism.";
RL   Am. J. Physiol. 284:E671-E678(2003).
RN   [10]
RP   REVIEW.
RX   PubMed=10322410;
RA   Drucker D.J.;
RT   "Glucagon-like peptide 2.";
RL   Trends Endocrinol. Metab. 10:153-156(1999).
RN   [11]
RP   REVIEW.
RX   MEDLINE=20073561; PubMed=10605628; DOI=10.1210/er.20.6.876;
RA   Kieffer T.J., Habener J.F.;
RT   "The glucagon-like peptides.";
RL   Endocr. Rev. 20:876-913(1999).
RN   [12]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 53-81.
RX   MEDLINE=76051297; PubMed=171582;
RA   Sasaki K., Dockerill S., Adamiak D.A., Tickle I.J., Blundell T.L.;
RT   "X-ray analysis of glucagon and its relationship to receptor
RT   binding.";
RL   Nature 257:751-757(1975).
CC   -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC       homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC       and decreasing glycolysis. A counterregulatory hormone of insulin,
CC       raises plasma glucose levels in response to insulin-induced
CC       hypoglycemia (By similarity).
CC   -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC       insulin release. Play important roles on gastric motility and the
CC       suppression of plasma glucagon levels. May be involved in the
CC       suppression of satiety and stimulation of glucose disposal in
CC       peripheral tissues, independent of the actions of insulin. Have
CC       growth-promoting activities on intestinal epithelium. May also
CC       regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC       TSH, CRH, oxytocin, and vasopressin (By similarity).
CC   -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC       villus height in the small intestine, concomitant with increased
CC       crypt cell proliferation and decreased enterocyte apoptosis. The
CC       gastrointestinal tract, from the stomach to the colon is the
CC       principal target for GLP-2 action. Plays a key role in nutrient
CC       homeostasis, enhancing nutrient assimilation through enhanced
CC       gastrointestinal function, as well as increasing nutrient
CC       disposal. Stimulates intestinal glucose transport and decreases
CC       mucosal permeability (By similarity).
CC   -!- FUNCTION: Oxyntomodulin significantly reduces food intake (By
CC       similarity).
CC   -!- FUNCTION: Glicentin may modulate gastric acid secretion and
CC       gastro-pyloro-duodenal activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Glucagon is secreted in the A cells of the
CC       islets of Langerhans. GLP-1, GLP-2, oxyntomodulin and glicentin
CC       are secreted from enteroendocrine cells throughout the
CC       gastrointestinal tract. GLP1 and GLP2 are also secreted in
CC       selected neurons in the brain.
CC   -!- INDUCTION: Glucagon release is stimulated by hypoglycemia and
CC       inhibited by hyperglycemia, insulin, and somatostatin. GLP-1 and
CC       GLP-2 are induced in response to nutrient ingestion (By
CC       similarity).
CC   -!- PTM: Proglucagon is posttranslationally processed in a tissue-
CC       specific manner in pancreatic A cells and intestinal L cells. In
CC       pancreatic A cells, the major bioactive hormone is glucagon
CC       cleaved by PCSK2/PC2. In the intestinal L cells PCSK1/PC1
CC       liberates GLP-1, GLP-2, glicentin and oxyntomodulin. GLP-1 is
CC       further N-terminally truncated by posttranslational processing in
CC       the intestinal L cells resulting in GLP-1(7-37) GLP-1-(7-36)amide.
```

```
CC       The C-terminal amidation is neither important for the metabolism
CC       of GLP-1 nor for its effects on the endocrine pancreas (By
CC       similarity).
CC   -!- MISCELLANEOUS: GLP-2 does not have cleavage on a pair of basic
CC       residues at C-terminus as in other mammmals.
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AY242124; AAO88211.1; -; mRNA.
DR   PDB; 1GCN; X-ray; @=53-81.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   3D-structure; Amidation; Cleavage on pair of basic residues;
KW   Direct protein sequencing; Glucagon family; Hormone; Signal.
FT   SIGNAL        1     20
FT   PEPTIDE      21     89       Glicentin.
FT   PEPTIDE      21     50       Glicentin-related polypeptide (By
FT                                similarity).
FT   PEPTIDE      53     89       Oxyntomodulin (By similarity).
FT   PEPTIDE      53     81       Glucagon.
FT   PROPEP       84     89       By similarity.
FT   PEPTIDE      92    128       Glucagon-like peptide 1 (By similarity).
FT   PEPTIDE      98    128       Glucagon-like peptide 1(7-37).
FT   PEPTIDE      98    127       Glucagon-like peptide 1(7-36).
FT   PROPEP      131    145       By similarity.
FT   PEPTIDE     146    180       Glucagon-like peptide 2 (By similarity).
FT   SITE         52     53       Cleavage (by PCSK2) (By similarity).
FT   SITE         83     84       Cleavage (by PCSK1 and PCSK2) (By
FT                                similarity).
FT   SITE         91     92       Cleavage (by PCSK1) (By similarity).
FT   SITE         97     98       Cleavage (by PCSK1) (By similarity).
FT   SITE        130    131       Cleavage (by PCSK1) (By similarity).
FT   SITE        145    146       Cleavage (by PCSK1) (By similarity).
FT   MOD_RES     127    127       Arginine amide (G-128 provides amide
FT                                group) (By similarity).
FT   CONFLICT    143    143       R -> G (in Ref. 6).
SQ   SEQUENCE   180 AA;  21029 MW;  362997AB72197EE6 CRC64;

  Query Match             52.6%;  Score 244.5;  DB 1;  Length 180;
  Best Local Similarity   42.7%;  Pred. No. 1.1e-17;
  Matches   53;  Conservative   15;  Mismatches   19;  Indels   37;  Gaps    3;

Qy          3 IFYIFLFLLSFVQG-----LEHTHRRG-----------------SLDKRHGEGTFTSDVS 40
              |:::    :  |||     |::|  :                  : |||| :|||||  |
Db          4 IYFVAGLFVMLVQGSWQRSLQNTEEKSRSFPAPQTDPLDDPDQMTEDKRHSQGTFTSDYS 63

Qy         41 SYLEGQAAKEFIAWLVK---------------GRHGEGTFTSDVSSYLEGQAAKEFIAWL 85
               ||: : |::|: ||:                || ||||||||||||||||||||||
Db         64 KYLDSRRAQDFVQWLMNTKRNKNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAWL 123

Qy         86 VKGR 89
              ||||
Db        124 VKGR 127


RESULT 6
GLUC_CANFA
ID   GLUC_CANFA     STANDARD;      PRT;   180 AA.
AC   P29794; Q95LG0;
DT   01-APR-1993 (Rel. 25, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE   (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE   (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE   peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)].
GN   Name=GCG;
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC   Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Pancreas, and Stomach;
RX   PubMed=11916259;
RA   Irwin D.M.;
RT   "cDNA cloning of proglucagon from the stomach and pancreas of the
RT   dog.";
RL   DNA Seq. 12:253-260(2001).
RN   [2]
RP   PROTEIN SEQUENCE OF 21-89.
RC   TISSUE=Ileum;
RX   MEDLINE=89185675; PubMed=3238052; DOI=10.1016/0167-0115(88)90230-3;
RA   Shinomura Y., Eng J., Yalow R.S.;
RT   "Immunoreactive glucagons purified from dog pancreas, stomach and
RT   ileum.";
RL   Regul. Pept. 23:299-308(1988).
RN   [3]
RP   PROCESSING BY PCSK1 AND PCSK2.
RX   PubMed=10499540; DOI=10.1210/en.140.10.4800;
RA   Damholt A.B., Buchan A.M., Holst J.J., Kofod H.;
RT   "Proglucagon processing profile in canine L cells expressing
RT   endogenous prohormone convertase 1/3 and prohormone convertase 2.";
RL   Endocrinology 140:4800-4808(1999).
RN   [4]
RP   REVIEW.
RX   PubMed=12554744; DOI=10.1210/me.2002-0306;
RA   Drucker D.J.;
RT   "Glucagon-like peptides: regulators of cell proliferation,
RT   differentiation, and apoptosis.";
RL   Mol. Endocrinol. 17:161-171(2003).
RN   [5]
RP   REVIEW.
RX   PubMed=12626323; DOI=10.1152/ajpendo.00492.2002;
RA   Jiang G., Zhang B.B.;
RT   "Glucagon and regulation of glucose metabolism.";
RL   Am. J. Physiol. 284:E671-E678(2003).
RN   [6]
RP   REVIEW.
RX   PubMed=10322410;
RA   Drucker D.J.;
RT   "Glucagon-like peptide 2.";
RL   Trends Endocrinol. Metab. 10:153-156(1999).
RN   [7]
RP   REVIEW.
RX   PubMed=10605628; DOI=10.1210/er.20.6.876;
RA   Kieffer T.J., Habener J.F.;
RT   "The glucagon-like peptides.";
RL   Endocr. Rev. 20:876-913(1999).
CC   -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC       homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC       and decreasing glycolysis. A counterregulatory hormone of insulin,
CC       raises plasma glucose levels in response to insulin-induced
CC       hypoglycemia (By similarity).
CC   -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC       insulin release. Play important roles on gastric motility and the
CC       suppression of plasma glucagon levels. May be involved in the
CC       suppression of satiety and stimulation of glucose disposal in
CC       peripheral tissues, independent of the actions of insulin. Have
CC       growth-promoting activities on intestinal epithelium. May also
CC       regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC       TSH, CRH, oxytocin, and vasopressin secretion. Increases islet
CC       mass through stimulation of islet neogenesis and pancreatic beta
CC       cell proliferaton (By similarity).
CC   -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC       villus height in the small intestine, concomitant with increased
CC       crypt cell proliferation and decreased enterocyte apoptosis. The
```

```
CC       gastrointestinal tract, from the stomach to the colon is the
CC       principal target for GLP-2 action. Plays a key role in nutrient
CC       homeostasis, enhancing nutrient assimilation through enhanced
CC       gastrointestinal function, as well as increasing nutrient
CC       disposal. Stimulates intestinal glucose transport and decreases
CC       mucosal permeability (By similarity).
CC   -!- FUNCTION: Oxyntomodulin significantly reduces food intake (By
CC       similarity).
CC   -!- FUNCTION: Glicentin may modulate gastric acid secretion and
CC       gastro-pyloro-duodenal activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Glucagon is secreted in the A cells of the
CC       islets of Langerhans. GLP-1, GLP-2, oxyntomodulin and glicentin
CC       are secreted from enteroendocrine cells throughout the
CC       gastrointestinal tract. GLP1 and GLP2 are also secreted in
CC       selected neurons in the brain.
CC   -!- INDUCTION: Glucagon release is stimulated by hypoglycemia and
CC       inhibited by hyperglycemia, insulin, and somatostatin. GLP-1 and
CC       GLP-2 are induced in response to nutrient ingestion (By
CC       similarity).
CC   -!- PTM: Proglucagon is posttranslationally processed in a tissue-
CC       specific manner in pancreatic A cells and intestinal L cells. In
CC       pancreatic A cells, the major bioactive hormone is glucagon
CC       cleaved by PCSK2/PC2. In the intestinal L cells PCSK1/PC1
CC       liberates GLP-1, GLP-2, glicentin and oxyntomodulin. GLP-1 is
CC       further N-terminally truncated by posttranslational processing in
CC       the intestinal L cells resulting in GLP-1(7-37) GLP-1-(7-36)amide.
CC       The C-terminal amidation is neither important for the metabolism
CC       of GLP-1 nor for its effects on the endocrine pancreas (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF308439; AAL09425.1; -; mRNA.
DR   PIR; A60318; GCDG69.
DR   HSSP; P01274; 1GCN.
DR   Ensembl; ENSCAFG00000010414; Canis familiaris.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Amidation; Cleavage on pair of basic residues;
KW   Direct protein sequencing; Glucagon family; Hormone; Signal.
FT   SIGNAL        1     20
FT   PEPTIDE      21     89       Glicentin.
FT   PEPTIDE      21     50       Glicentin-related polypeptide (By
FT                                similarity).
FT   PEPTIDE      53     89       Oxyntomodulin (By similarity).
FT   PEPTIDE      53     81       Glucagon (By similarity).
FT   PROPEP       84     89       By similarity.
FT   PEPTIDE      92    128       Glucagon-like peptide 1.
FT   PEPTIDE      98    128       Glucagon-like peptide 1(7-37).
FT   PEPTIDE      98    127       Glucagon-like peptide 1(7-36).
FT   PROPEP      131    145       By similarity.
FT   PEPTIDE     146    178       Glucagon-like peptide 2 (By similarity).
FT   SITE         52     53       Cleavage (by PCSK2).
FT   SITE         83     84       Cleavage (by PCSK1 and PCSK2).
FT   SITE         91     92       Cleavage (by PCSK1).
FT   SITE         97     98       Cleavage (by PCSK1).
FT   SITE        130    131       Cleavage (by PCSK1).
FT   SITE        145    146       Cleavage (by PCSK1).
FT   MOD_RES     127    127       Arginine amide (G-128 provides amide
FT                                group) (By similarity).
SQ   SEQUENCE   180 AA;  21115 MW;  80F66941AFC324FD CRC64;

  Query Match             52.4%;  Score 243.5;  DB 1;  Length 180;
  Best Local Similarity   42.4%;  Pred. No. 1.4e-17;

  Matches   53;  Conservative   15;  Mismatches   20;  Indels   37;  Gaps    3;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRG-----------------SLDKRHGEGTFTSDV 39
              :|:::    :  |||     |: |  :                   : |||| :||||||
Db          3 SIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSAPQTEPLNDLDQMNEDKRHSQGTFTSDY 62

Qy         40 SSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSSYLEGQAAKEFIAW 84
              | ||: : |::|: ||:                || ||||||||||||||||||||||
Db         63 SKYLDSRRAQDFVQWLMNTKRNKNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAW 122

Qy         85 LVKGR 89
              |||||
Db        123 LVKGR 127


RESULT 7
GLUC_MOUSE
ID   GLUC_MOUSE     STANDARD;      PRT;   180 AA.
AC   P55095;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE   (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE   (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE   peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)].
GN   Name=Gcg;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Pancreatic islets;
RX   MEDLINE=95247722; PubMed=7730317; DOI=10.1074/jbc.270.17.10136;
RA   Rothenberg M.E., Eilertson C.D., Klein K., Zhou Y., Linberg I.,
RA   McDonald J.K., Mackin R.B., Noe B.D.;
RT   "Processing of mouse proglucagon by recombinant prohormone convertase
RT   1 and immunopurified prohormone convertase 2 in vitro.";
RL   J. Biol. Chem. 270:10136-10146(1995).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Shamsadin R., Knepel W.;
RT   "Mouse glucagon full length cDNA.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=C57BL/6J; TISSUE=Pancreas;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
```

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP FUNCTION OF GLP-1 AND GLP-1(7-36) AMIDE.
RX PubMed=1886889;
RA Fridolf T., Bottcher G., Sundler F., Ahren B.;
RT "GLP-1 and GLP-1(7-36) amide: influences on basal and stimulated
RT insulin and glucagon secretion in the mouse.";
RL Pancreas 6:208-215(1991).
RN [6]
RP PROCESSING BY PCSK1.
RX PubMed=9407057; DOI=10.1074/jbc.272.52.32810;
RA Rouille Y., Kantengwa S., Irminger J.C., Halban P.A.;
RT "Role of the prohormone convertase PC3 in the processing of
RT proglucagon to glucagon-like peptide 1.";
RL J. Biol. Chem. 272:32810-32816(1997).
RN [7]
RP PROCESSING BY PCSK2.
RX PubMed=11356850; DOI=10.1074/jbc.M103362200;
RA Furuta M., Zhou A., Webb G., Carroll R., Ravazzola M., Orci L.,
RA Steiner D.F.;
RT "Severe defect in proglucagon processing in islet A-cells of
RT prohormone convertase 2 null mice.";
RL J. Biol. Chem. 276:27197-27202(2001).
RN [8]
RP REVIEW.
RX PubMed=12554744; DOI=10.1210/me.2002-0306;
RA Drucker D.J.;
RT "Glucagon-like peptides: regulators of cell proliferation,
RT differentiation, and apoptosis.";
RL Mol. Endocrinol. 17:161-171(2003).
RN [9]
RP REVIEW.
RX PubMed=12626323; DOI=10.1152/ajpendo.00492.2002;
RA Jiang G., Zhang B.B.;
RT "Glucagon and regulation of glucose metabolism.";
RL Am. J. Physiol. 284:E671-E678(2003).
RN [10]
RP REVIEW.
RX PubMed=10322410;
RA Drucker D.J.;
RT "Glucagon-like peptide 2.";
RL Trends Endocrinol. Metab. 10:153-156(1999).
RN [11]
RP REVIEW.
RX PubMed=10605628; DOI=10.1210/er.20.6.876;
RA Kieffer T.J., Habener J.F.;
RT "The glucagon-like peptides.";
RL Endocr. Rev. 20:876-913(1999).
CC -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC and decreasing glycolysis. A counterregulatory hormone of insulin,
CC raises plasma glucose levels in response to insulin-induced
CC hypoglycemia (By similarity).
CC -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC insulin release. Play important roles on gastric motility and the
CC suppression of plasma glucagon levels. May be involved in the
CC suppression of satiety and stimulation of glucose disposal in
CC peripheral tissues, independent of the actions of insulin. Have
CC growth-promoting activities on intestinal epithelium. May also
CC regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC TSH, CRH, oxytocin, and vasopressin secretion. Increases islet
CC mass through stimulation of islet neogenesis and pancreatic beta
CC cell proliferaton (By similarity).
CC -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC villus height in the small intestine, concomitant with increased
CC crypt cell proliferation and decreased enterocyte apoptosis. The
CC gastrointestinal tract, from the stomach to the colon is the
CC principal target for GLP-2 action. Plays a key role in nutrient
CC homeostasis, enhancing nutrient assimilation through enhanced
CC gastrointestinal function, as well as increasing nutrient
CC disposal. Stimulates intestinal glucose transport and decreases
CC mucosal permeability (By similarity).
CC -!- FUNCTION: Oxyntomodulin significantly reduces food intake (By
CC similarity).
CC -!- FUNCTION: Glicentin may modulate gastric acid secretion and
CC gastro-pyloro-duodenal activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Glucagon is secreted in the A cells of the
CC islets of Langerhans. GLP-1, GLP-2, oxyntomodulin and glicentin
CC are secreted from enteroendocrine cells throughout the
CC gastrointestinal tract. GLP1 and GLP2 are also secreted in
CC selected neurons in the brain.
CC -!- INDUCTION: Glucagon release is stimulated by hypoglycemia and
CC inhibited by hyperglycemia, insulin, and somatostatin. GLP-1 and
CC GLP-2 are induced in response to nutrient ingestion (By
CC similarity).
CC -!- PTM: Proglucagon is posttranslationally processed in a tissue-
CC specific manner in pancreatic A cells and intestinal L cells. In
CC pancreatic A cells, the major bioactive hormone is glucagon
CC cleaved by PCSK2/PC2. In the intestinal L cells PCSK1/PC1
CC liberates GLP-1, GLP-2, glicentin and oxyntomodulin. GLP-1 is
CC further N-terminally truncated by posttranslational processing in
CC the intestinal L cells resulting in GLP-1(7-37) GLP-1-(7-36)amide.
CC The C-terminal amidation is neither important for the metabolism
CC of GLP-1 nor for its effects on the endocrine pancreas (By
CC similarity).
CC -!- SIMILARITY: Belongs to the glucagon family.
CC ---

CC ---
DR EMBL; Z46845; CAA86902.1; -; mRNA.
DR EMBL; AF276754; AAK96898.1; -; mRNA.
DR EMBL; AK007911; BAB25344.1; -; mRNA.
DR EMBL; BC012975; AAH12975.1; -; mRNA.
DR PIR; A57294; A57294.
DR HSSP; P01275; 1D0R.
DR Ensembl; ENSMUSG00000000394; Mus musculus.
DR MGI; MGI:95674; Gcg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000532; Glucagon.
DR Pfam; PF00123; Hormone_2; 3.

Query Match 52.4%; Score 243.5; DB 1; Length 180;

```
  Best Local Similarity   43.5%;  Pred. No. 1.4e-17;
  Matches   54;  Conservative   13;  Mismatches   20;  Indels   37;  Gaps    3;

Qy          3 IFYIFLFLLSFVQG-----LEHTHRRG----------------SLDKRHGEGTFTSDVS 40
              |:::   |:  |||      |: |                    : |||| :|||||| |
Db          4 IYFVAGLLIMLVQGSWQHALQDTEENPRSFPASQTEAHEDPDEMNEDKRHSQGTFTSDYS 63

Qy         41 SYLEGQAAKEFIAWLVK---------------GRHGEGTFTSDVSSYLEGQAAKEFIAWL 85
               ||: : |::|: ||:                || ||||||||||||||||||||||
Db         64 KYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAWL 123

Qy         86 VKGR 89
              ||||
Db        124 VKGR 127


RESULT 8
Q53TP6_HUMAN
ID   Q53TP6_HUMAN PRELIMINARY;      PRT;   180 AA.
AC   Q53TP6;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein GCG.
GN   Name=GCG;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Cotton M., Maupin R., Hawkins M., Harkins R.;
RT   "The sequence of Homo sapiens BAC clone RP11-576I16.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Waterston R.H.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Waterston R.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RA   Wilson R.K.;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007750; AAY24204.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   180 AA;  20909 MW;  7A99EEC629B2862C CRC64;

  Query Match             52.4%;  Score 243.5;  DB 2;  Length 180;
  Best Local Similarity   42.4%;  Pred. No. 1.4e-17;
  Matches   53;  Conservative   15;  Mismatches   20;  Indels   37;  Gaps    3;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRG----------------SLDKRHGEGTFTSDV 39
              :|:::    :  |||      |: |  :                : |||| :||||||
Db          3 SIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDPDQMNEDKRHSQGTFTSDY 62

Qy         40 SSYLEGQAAKEFIAWLVK---------------GRHGEGTFTSDVSSYLEGQAAKEFIAW 84
              | ||: : |::|: ||:                || ||||||||||||||||||||||
Db         63 SKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAW 122

Qy         85 LVKGR 89
              |||||
Db        123 LVKGR 127


RESULT 9
GLUC_HUMAN
ID   GLUC_HUMAN     STANDARD;      PRT;   180 AA.
AC   P01275;
DT   21-JUL-1986 (Rel. 01, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE   (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE   (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE   peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)].
GN   Name=GCG;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=88330860; PubMed=2901414;
RA   Drucker D.J., Asa S.;
RT   "Glucagon gene expression in vertebrate brain.";
RL   J. Biol. Chem. 263:13475-13478(1988).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=86259053; PubMed=3725587;
RA   White J.W., Saunders G.F.;
RT   "Structure of the human glucagon gene.";
RL   Nucleic Acids Res. 14:4719-4730(1986).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=83271477; PubMed=6877358;
RA   Bell G.I., Sanchez-Pescador R., Laybourn P.J., Najarian R.C.;
RT   "Exon duplication and divergence in the human preproglucagon gene.";
RL   Nature 304:368-371(1983).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RA   Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA   Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA   Phelan M., Farmer A.;
RT   "Cloning of human full-length CDSs in BD Creator(TM) system donor
RT   vector.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Pancreas;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   PROTEIN SEQUENCE OF 53-81.
RX   PubMed=11946536;
RA   Thomsen J., Kristiansen K., Brunfeldt K., Sundby F.;
RT   "The amino acid sequence of human glucagon.";
RL   FEBS Lett. 21:315-319(1972).
RN   [7]
RP   PROTEIN SEQUENCE OF 98-127.
RX   MEDLINE=89327238; PubMed=2753890;
RA   Orskov C., Bersani M., Johnsen A.H., Hoejrup P., Holst J.J.;
```

```
RT   "Complete sequences of glucagon-like peptide-1 from human and pig
RT   small intestine.";
RL   J. Biol. Chem. 264:12826-12829(1989).
RN   [8]
RP   FUNCTION OF GLP1 BIOACTIVE FORMS.
RX   MEDLINE=93246081; PubMed=8482423;
RA   Orskov C., Wettergren A., Holst J.J.;
RT   "Biological effects and metabolic rates of glucagonlike peptide-1 7-36
RT   amide and glucagonlike peptide-1 7-37 in healthy subjects are
RT   indistinguishable.";
RL   Diabetes 42:658-661(1993).
RN   [9]
RP   FUNCTION OF OXYNTOMODULIN.
RX   MEDLINE=22919492; PubMed=14557443; DOI=10.1210/jc.2003-030421;
RA   Cohen M.A., Ellis S.M., Le Roux C.W., Batterham R.L., Park A.,
RA   Patterson M., Frost G.S., Ghatei M.A., Bloom S.R.;
RT   "Oxyntomodulin suppresses appetite and reduces food intake in
RT   humans.";
RL   J. Clin. Endocrinol. Metab. 88:4696-4701(2003).
RN   [10]
RP   FUNCTION OF GLICENTIN.
RX   MEDLINE=22993785; PubMed=14632334; DOI=10.1080/08035250310000514;
RA   Tadokoro R., Shimizu T., Hosaka A., Kaneko N., Satoh Y., Yamashiro Y.;
RT   "Postnatal and postprandial changes in plasma concentrations of
RT   glicentin in term and preterm infants.";
RL   Acta Paediatr. 92:1175-1179(2003).
RN   [11]
RP   PROCESSING BY PCSK2.
RX   MEDLINE=97431623; PubMed=9287128; DOI=10.1016/S0014-5793(97)00892-2;
RA   Rouille Y., Bianchi M., Irminger J.C., Halban P.A.;
RT   "Role of the prohormone convertase PC2 in the processing of
RT   proglucagon to glucagon.";
RL   FEBS Lett. 413:119-123(1997).
RN   [12]
RP   PROCESSING BY PCSK1.
RX   MEDLINE=22538931; PubMed=12651102; DOI=10.1016/S1046-5928(02)00653-8;
RA   Bonic A., Mackin R.B.;
RT   "Expression, purification, and PC1-mediated processing of human
RT   proglucagon, glicentin, and major proglucagon fragment.";
RL   Protein Expr. Purif. 28:15-24(2003).
RN   [13]
RP   REVIEW.
RX   PubMed=14719035; DOI=10.1139/y03-107;
RA   Brubaker P.L., Anini Y.;
RT   "Direct and indirect mechanisms regulating secretion of glucagon-like
RT   peptide-1 and glucagon-like peptide-2.";
RL   Can. J. Physiol. Pharmacol. 81:1005-1012(2003).
RN   [14]
RP   REVIEW.
RX   MEDLINE=22442611; PubMed=12554744; DOI=10.1210/me.2002-0306;
RA   Drucker D.J.;
RT   "Glucagon-like peptides: regulators of cell proliferation,
RT   differentiation, and apoptosis.";
RL   Mol. Endocrinol. 17:161-171(2003).
RN   [15]
RP   REVIEW.
RX   MEDLINE=22513095; PubMed=12626323; DOI=10.1152/ajpendo.00492.2002;
RA   Jiang G., Zhang B.B.;
RT   "Glucagon and regulation of glucose metabolism.";
RL   Am. J. Physiol. 284:E671-E678(2003).
RN   [16]
RP   REVIEW.
RX   PubMed=10322410;
RA   Drucker D.J.;
RT   "Glucagon-like peptide 2.";
RL   Trends Endocrinol. Metab. 10:153-156(1999).
RN   [17]
RP   REVIEW.
RX   MEDLINE=20073561; PubMed=10605628; DOI=10.1210/er.20.6.876;
RA   Kieffer T.J., Habener J.F.;
RT   "The glucagon-like peptides.";
RL   Endocr. Rev. 20:876-913(1999).
RN   [18]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 53-81.
RX   MEDLINE=98334683; PubMed=9667960; DOI=10.1021/jm980084a;
RA   Sturm N.S., Lin Y., Burley S.K., Krstenansky J.L., Ahn J.-M.,
RA   Azizeh B.Y., Trivedi D., Hruby V.J.;
RT   "Structure-function studies on positions 17, 18, and 21 replacement
RT   analogues of glucagon: the importance of charged residues and salt
RT   bridges in glucagon biological activity.";
RL   J. Med. Chem. 41:2693-2700(1998).
RN   [19]
RP   STRUCTURE BY NMR OF 98-127.
RX   MEDLINE=21940600; PubMed=11943215; DOI=10.1016/S0014-5793(02)02466-3;
RA   Chang X., Keller D., O'Donoghue S.I., Led J.J.;
RT   "NMR studies of the aggregation of glucagon-like peptide-1: formation
RT   of a symmetric helical dimer.";
RL   FEBS Lett. 515:165-170(2002).
RN   [20]
RP   STRUCTURE BY NMR OF GLUCAGON ANTAGONIST.
RX   MEDLINE=22515400; PubMed=12627948; DOI=10.1021/bi026629r;
RA   Ying J., Ahn J.-M., Jacobsen N.E., Brown M.F., Hruby V.J.;
RT   "NMR solution structure of the glucagon antagonist [desHis1, desPhe6,
RT   Glu9]glucagon amide in the presence of perdeuterated
RT   dodecylphosphocholine micelles.";
RL   Biochemistry 42:2825-2835(2003).
CC   -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC       homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC       and decreasing glycolysis. A counterregulatory hormone of insulin,
CC       raises plasma glucose levels in response to insulin-induced
CC       hypoglycemia. Plays an important role in initiating and
CC       maintaining hyperglycemic conditions in diabetes.
CC   -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC       insulin release. Play important roles on gastric motility and the
CC       suppression of plasma glucagon levels. May be involved in the
CC       suppression of satiety and stimulation of glucose disposal in
CC       peripheral tissues, independent of the actions of insulin. Have
CC       growth-promoting activities on intestinal epithelium. May also
CC       regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC       TSH, CRH, oxytocin, and vasopressin secretion. Increases islet
CC       mass through stimulation of islet neogenesis and pancreatic beta
CC       cell proliferaton. Inhibits beta cell apoptosis.
CC   -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC       villus height in the small intestine, concomitant with increased
CC       crypt cell proliferation and decreased enterocyte apoptosis. The
CC       gastrointestinal tract, from the stomach to the colon is the
CC       principal target for GLP-2 action. Plays a key role in nutrient
CC       homeostasis, enhancing nutrient assimilation through enhanced
CC       gastrointestinal function, as well as increasing nutrient
CC       disposal. Stimulates intestinal glucose transport and decreases
CC       mucosal permeability.
CC   -!- FUNCTION: Oxyntomodulin significantly reduces food intake.
CC       Inhibits gastric emptying in humans. Suppression of gastric
CC       emptying may lead to increased gastric distension, which may

  Query Match             51.9%;  Score 241.5;  DB 1;  Length 180;
  Best Local Similarity   60.3%;  Pred. No. 2.3e-17;
  Matches   47;  Conservative    7;  Mismatches    9;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSS 71
              |||| :|||||| | ||: : |::|: ||:                || ||||||||||
Db         50 DKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


RESULT 10
GLUC_RAT
ID   GLUC_RAT       STANDARD;      PRT;   180 AA.
AC   P06883;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
```

```
DE   (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE   (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE   peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)].
GN   Name=Gcg;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=85054853; PubMed=6094539;
RA   Heinrich G., Gros P., Habener J.F.;
RT   "Glucagon gene sequence. Four of six exons encode separate functional
RT   domains of rat pre-proglucagon.";
RL   J. Biol. Chem. 259:14082-14087(1984).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=85051023; PubMed=6548696;
RA   Heinrich G., Gros P., Lund P.K., Bentley R.C., Habener J.F.;
RT   "Pre-proglucagon messenger ribonucleic acid: nucleotide and encoded
RT   amino acid sequences of the rat pancreatic complementary
RT   deoxyribonucleic acid.";
RL   Endocrinology 115:2176-2181(1984).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=86304324; PubMed=3528148;
RA   Mojsov S., Heinrich G., Wilson I.B., Ravazzola M., Orci L.,
RA   Habener J.F.;
RT   "Preproglucagon gene expression in pancreas and intestine diversifies
RT   at the level of post-translational processing.";
RL   J. Biol. Chem. 261:11880-11889(1986).
RN   [4]
RP   PROTEIN SEQUENCE OF 53-89.
RX   MEDLINE=95023911; PubMed=7937770;
RA   Collie N.L., Walsh J.H., Wong H.C., Shively J.E., Davis M.T.,
RA   Lee T.D., Reeve J.R. Jr.;
RT   "Purification and sequence of rat oxyntomodulin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:9362-9366(1994).
RN   [5]
RP   FUNCTION OF OXYNTOMODULIN.
RX   MEDLINE=21448403; PubMed=11564680; DOI=10.1210/en.142.10.4244;
RA   Dakin C.L., Gunn I., Small C.J., Edwards C.M., Hay D.L., Smith D.M.,
RA   Ghatei M.A., Bloom S.R.;
RT   "Oxyntomodulin inhibits food intake in the rat.";
RL   Endocrinology 142:4244-4250(2001).
RN   [6]
RP   PROCESSING BY PCSK1 AND PCSK2.
RX   MEDLINE=96282838; PubMed=8721980; DOI=10.1210/me.10.4.342;
RA   Dhanvantari S., Seidah N.G., Brubaker P.L.;
RT   "Role of prohormone convertases in the tissue-specific processing of
RT   proglucagon.";
RL   Mol. Endocrinol. 10:342-355(1996).
RN   [7]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=90243673; PubMed=1692320;
RA   Mojsov S., Kopczynski M.G., Habener J.F.;
RT   "Both amidated and nonamidated forms of glucagon-like peptide I are
RT   synthesized in the rat intestine and the pancreas.";
RL   J. Biol. Chem. 265:8001-8008(1990).
RN   [8]
RP   REVIEW.
RX   PubMed=14719035; DOI=10.1139/y03-107;
RA   Brubaker P.L., Anini Y.;
RT   "Direct and indirect mechanisms regulating secretion of glucagon-like
RT   peptide-1 and glucagon-like peptide-2.";
RL   Can. J. Physiol. Pharmacol. 81:1005-1012(2003).
RN   [9]
RP   REVIEW.
RX   MEDLINE=22442611; PubMed=12554744; DOI=10.1210/me.2002-0306;
RA   Drucker D.J.;
RT   "Glucagon-like peptides: regulators of cell proliferation,
RT   differentiation, and apoptosis.";
RL   Mol. Endocrinol. 17:161-171(2003).
RN   [10]
RP   REVIEW.
RX   MEDLINE=22513095; PubMed=12626323; DOI=10.1152/ajpendo.00492.2002;
RA   Jiang G., Zhang B.B.;
RT   "Glucagon and regulation of glucose metabolism.";
RL   Am. J. Physiol. 284:E671-E678(2003).
RN   [11]
RP   REVIEW.
RX   PubMed=10322410;
RA   Drucker D.J.;
RT   "Glucagon-like peptide 2.";
RL   Trends Endocrinol. Metab. 10:153-156(1999).
RN   [12]
RP   REVIEW.
RX   MEDLINE=20073561; PubMed=10605628; DOI=10.1210/er.20.6.876;
RA   Kieffer T.J., Habener J.F.;
RT   "The glucagon-like peptides.";
RL   Endocr. Rev. 20:876-913(1999).
CC   -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC       homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC       and decreasing glycolysis. A counterregulatory hormone of insulin,
CC       raises plasma glucose levels in response to insulin-induced
CC       hypoglycemia. Plays an important role in initiating and
CC       maintaining hyperglycemic conditions in diabetes.
CC   -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC       insulin release. Play important roles on gastric motility and the
CC       suppression of plasma glucagon levels. May be involved in the
CC       suppression of satiety and stimulation of glucose disposal in
CC       peripheral tissues, independent of the actions of insulin. Have
CC       growth-promoting activities on intestinal epithelium. May also
CC       regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC       TSH, CRH, oxytocin, and vasopressin secretion. Increases islet
CC       mass through stimulation of islet neogenesis and pancreatic beta
CC       cell proliferaton. Inhibits beta cell apoptosis.
CC   -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC       villus height in the small intestine, concomitant with increased
CC       crypt cell proliferation and decreased enterocyte apoptosis. The
CC       gastrointestinal tract, from the stomach to the colon is the
CC       principal target for GLP-2 action. Plays a key role in nutrient
CC       homeostasis, enhancing nutrient assimilation through enhanced
CC       gastrointestinal function, as well as increasing nutrient
CC       disposal. Stimulates intestinal glucose transport and decreases
CC       mucosal permeability.
CC   -!- FUNCTION: Oxyntomodulin significantly reduces food intake.
CC   -!- FUNCTION: Glicentin may modulate gastric acid secretion and the
CC       gastro-pyloro-duodenal activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Glucagon is secreted in the A cells of the
CC       islets of Langerhans. GLP-1, GLP-2, oxyntomodulin and glicentin
CC       are secreted from enteroendocrine cells throughout the
CC       gastrointestinal tract.
CC   -!- INDUCTION: Glucagon release is stimulated by hypoglycemia and
CC       inhibited by hyperglycemia, insulin, and somatostatin. GLP-1 and
CC       GLP-2 are induced in response to nutrient ingestion.
CC   -!- PTM: Proglucagon is posttranslationally processed in a tissue-
CC       specific manner in pancreatic A cells and intestinal L cells. In
CC       pancreatic A cells, the major bioactive hormone is glucagon
CC       cleaved by PCSK2/PC2. In the intestinal L cells PCSK1/PC1
CC       liberates GLP-1, GLP-2, glicentin and oxyntomodulin. GLP-1 is
CC       further N-terminally truncated by posttranslational processing in
CC       the intestinal L cells resulting in GLP-1(7-37) GLP-1-(7-36)amide.
CC       The C-terminal amidation is neither important for the metabolism
CC       of GLP-1 nor for its effects on the endocrine pancreas.
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; K02813; AAA41235.1; -; Genomic_DNA.
```

```
DR   EMBL; K02809; AAA41235.1; JOINED; Genomic_DNA.
DR   EMBL; K02810; AAA41235.1; JOINED; Genomic_DNA.
DR   EMBL; K02811; AAA41235.1; JOINED; Genomic_DNA.
DR   EMBL; K02812; AAA41235.1; JOINED; Genomic_DNA.
DR   PIR; A22655; GCRT.
DR   HSSP; P01275; 1D0R.
DR   Ensembl; ENSRNOG00000005498; Rattus norvegicus.
DR   RGD; 2668; Gcg.
DR   GO; GO:0005179; F:hormone activity; TAS.
DR   GO; GO:0019538; P:protein metabolism; TAS.
DR   GO; GO:0006109; P:regulation of carbohydrate metabolism; TAS.
DR   GO; GO:0019216; P:regulation of lipid metabolism; TAS.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Amidation; Cleavage on pair of basic residues;
KW   Direct protein sequencing; Glucagon family; Hormone; Signal.
FT   SIGNAL        1     20
FT   PEPTIDE      21     89       Glicentin (By similarity).
FT   PEPTIDE      21     50       Glicentin-related polypeptide (By
FT                                similarity).
FT   PEPTIDE      53     89       Oxyntomodulin.
FT   PEPTIDE      53     81       Glucagon (By similarity).
FT   PROPEP       84     89       By similarity.
FT   PEPTIDE      92    128       Glucagon-like peptide 1 (By similarity).
FT   PEPTIDE      98    128       Glucagon-like peptide 1(7-37) (By
FT                                similarity).
FT   PEPTIDE      98    127       Glucagon-like peptide 1(7-36) (By
FT                                similarity).
FT   PROPEP      131    145       By similarity.
FT   PEPTIDE     146    178       Glucagon-like peptide 2 (By similarity).
FT   SITE         52     53       Cleavage (by PCSK2) (By similarity).
FT   SITE         83     84       Cleavage (by PCSK1 and PCSK2) (By
FT                                similarity).
FT   SITE         91     92       Cleavage (by PCSK1) (By similarity).
FT   SITE         97     98       Cleavage (by PCSK1) (By similarity).
FT   SITE        130    131       Cleavage (by PCSK1) (By similarity).
FT   SITE        145    146       Cleavage (by PCSK1) (By similarity).
FT   MOD_RES     127    127       Arginine amide (G-128 provides amide
FT                                group) (By similarity).
SQ   SEQUENCE   180 AA;  20846 MW;  76931409D03C7978 CRC64;

  Query Match             51.9%;  Score 241.5;  DB 1;  Length 180;
  Best Local Similarity   60.3%;  Pred. No. 2.3e-17;
  Matches   47;  Conservative    7;  Mismatches    9;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK---------------GRHGEGTFTSDVSS 71
              |||| :|||||| | ||: : |::|: ||:                 || ||||||||||
Db         50 DKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


RESULT 11
GLUC_CAVPO
ID   GLUC_CAVPO     STANDARD;      PRT;   180 AA.
AC   P05110;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE   (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE   (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE   peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)].
GN   Name=GCG;
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
OC   Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=86248118; PubMed=3755107; DOI=10.1016/0014-5793(86)81429-6;
RA   Seino S., Welsh M., Bell G.I., Chan S.J., Steiner D.F.;
RT   "Mutations in the guinea pig preproglucagon gene are restricted to a
RT   specific portion of the prohormone sequence.";
RL   FEBS Lett. 203:25-30(1986).
RN   [2]
RP   PROTEIN SEQUENCE OF 53-81.
RX   MEDLINE=86165412; PubMed=3956884;
RA   Huang C.G., Eng J., Pan Y.-C.E., Hulmes J.D., Yalow R.S.;
RT   "Guinea pig glucagon differs from other mammalian glucagons.";
RL   Diabetes 35:508-512(1986).
RN   [3]
RP   PARTIAL PROTEIN SEQUENCE OF 53-89.
RX   MEDLINE=86017849; PubMed=4048553; DOI=10.1016/0167-0115(85)90203-4;
RA   Conlon J.M., Hansen H.F., Schwartz T.W.;
RT   "Primary structure of glucagon and a partial sequence of oxyntomodulin
RT   (glucagon-37) from the guinea pig.";
RL   Regul. Pept. 11:309-320(1985).
RN   [4]
RP   REVIEW.
RX   MEDLINE=22442611; PubMed=12554744; DOI=10.1210/me.2002-0306;
RA   Drucker D.J.;
RT   "Glucagon-like peptides: regulators of cell proliferation,
RT   differentiation, and apoptosis.";
RL   Mol. Endocrinol. 17:161-171(2003).
RN   [5]
RP   REVIEW.
RX   MEDLINE=22513095; PubMed=12626323; DOI=10.1152/ajpendo.00492.2002;
RA   Jiang G., Zhang B.B.;
RT   "Glucagon and regulation of glucose metabolism.";
RL   Am. J. Physiol. 284:E671-E678(2003).
RN   [6]
RP   REVIEW.
RX   PubMed=10322410;
RA   Drucker D.J.;
RT   "Glucagon-like peptide 2.";
RL   Trends Endocrinol. Metab. 10:153-156(1999).
RN   [7]
RP   REVIEW.
RX   MEDLINE=20073561; PubMed=10605628; DOI=10.1210/er.20.6.876;
RA   Kieffer T.J., Habener J.F.;
RT   "The glucagon-like peptides.";
RL   Endocr. Rev. 20:876-913(1999).
CC   -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC       homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC       and decreasing glycolysis. A counterregulatory hormone of insulin,
CC       raises plasma glucose levels in response to insulin-induced
CC       hypoglycemia (By similarity).
CC   -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC       insulin release. Play important roles on gastric motility and the
CC       suppression of plasma glucagon levels. May be involved in the
CC       suppression of satiety and stimulation of glucose disposal in
CC       peripheral tissues, independent of the actions of insulin. Have
CC       growth-promoting activities on intestinal epithelium. May also
CC       regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC       TSH, CRH, oxytocin, and vasopressin secretion. Increases islet
CC       mass through stimulation of islet neogenesis and pancreatic beta
CC       cell proliferaton (By similarity).
CC   -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC       villus height in the small intestine, concomitant with increased
CC       crypt cell proliferation and decreased enterocyte apoptosis. The
CC       gastrointestinal tract, from the stomach to the colon is the
CC       principal target for GLP-2 action. Plays a key role in nutrient
CC       homeostasis, enhancing nutrient assimilation through enhanced
CC       gastrointestinal function, as well as increasing nutrient
CC       disposal. Stimulates intestinal glucose transport and decreases
CC       mucosal permeability (By similarity).
CC   -!- FUNCTION: Oxyntomodulin significantly reduces food intake (By
CC       similarity).
CC   -!- FUNCTION: Glicentin may modulate gastric acid secretion and
CC       gastro-pyloro-duodenal activity (By similarity).
```

```
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- INDUCTION: Glucagon release is stimulated by hypoglycemia and
CC       inhibited by hyperglycemia, insulin, and somatostatin. GLP-1 and
CC       GLP-2 are induced in response to nutrient ingestion (By
CC       similarity).
CC   -!- PTM: Proglucagon is posttranslationally processed in a tissue-
CC       specific manner in pancreatic A cells and intestinal L cells. In
CC       pancreatic A cells, the major bioactive hormone is glucagon
CC       cleaved by PCSK2/PC2. In the intestinal L cells PCSK1/PC1
CC       liberates GLP-1, GLP-2, glicentin and oxyntomodulin. GLP-1 is
CC       further N-terminally truncated by posttranslational processing in
CC       the intestinal L cells resulting in GLP-1(7-37) GLP-1-(7-36)amide.
CC       The C-terminal amidation is neither important for the metabolism
CC       of GLP-1 nor for its effects on the endocrine pancreas (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; D00014; BAA00010.1; -; mRNA.
DR   PIR; A24856; GCGP.
DR   HSSP; P01275; 1D0R.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Amidation; Cleavage on pair of basic residues;
KW   Direct protein sequencing; Glucagon family; Hormone; Signal.
FT   SIGNAL        1     20
FT   PEPTIDE      21     89       Glicentin (By similarity).
FT   PEPTIDE      21     50       Glicentin-related polypeptide (By
FT                                similarity).
FT   PEPTIDE      53     89       Oxyntomodulin.
FT   PEPTIDE      53     81       Glucagon.
FT   PROPEP       84     89       By similarity.
FT   PEPTIDE      92    128       Glucagon-like peptide 1 (By similarity).
FT   PEPTIDE      98    128       Glucagon-like peptide 1(7-37) (By
FT                                similarity).
FT   PEPTIDE      98    127       Glucagon-like peptide 1(7-36) (By
FT                                similarity).
FT   PROPEP      131    145       By similarity.
FT   PEPTIDE     146    178       Glucagon-like peptide 2 (By similarity).
FT   SITE         52     53       Cleavage (by PCSK2) (By similarity).
FT   SITE         83     84       Cleavage (by PCSK1 and PCSK2) (By
FT                                similarity).
FT   SITE         91     92       Cleavage (by PCSK1) (By similarity).
FT   SITE         97     98       Cleavage (by PCSK1) (By similarity).
FT   SITE        130    131       Cleavage (by PCSK1) (By similarity).
FT   SITE        145    146       Cleavage (by PCSK1) (By similarity).
FT   MOD_RES     127    127       Arginine amide (G-128 provides amide
FT                                group) (By similarity).
SQ   SEQUENCE   180 AA;  20972 MW;  702FB181161D2776 CRC64;

  Query Match             51.7%;  Score 240.5;  DB 1;  Length 180;
  Best Local Similarity   60.3%;  Pred. No. 3e-17;
  Matches   47;  Conservative    7;  Mismatches    9;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSS 71
              |||| :|||||| | ||: : |::|: ||:              || ||||||||||
Db         50 DKRHSQGTFTSDYSKYLDSRRAQQFLKWLLNVKRNRNNIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


RESULT 12
Q6RYB2_BUFMA
ID   Q6RYB2_BUFMA PRELIMINARY;      PRT;   149 AA.
AC   Q6RYB2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Proglucagon (Fragment).
OS   Bufo marinus (Giant toad) (Cane toad).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Bufonidae; Bufo.
OX   NCBI_TaxID=8386;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Intestine;
RA   Busby E.R., Brown G.D., Mommsen T.P.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY485819; AAS57655.1; -; mRNA.
DR   GO; GO:0005576; C:extracellular region; IEA.
DR   GO; GO:0005179; F:hormone activity; IEA.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 3.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 3.
FT   NON_TER       1      1
SQ   SEQUENCE   149 AA;  17322 MW;  2F99199A0778B8AF CRC64;

  Query Match             51.2%;  Score 238;  DB 2;  Length 149;
  Best Local Similarity   51.6%;  Pred. No. 4.4e-17;
  Matches   47;  Conservative   16;  Mismatches   14;  Indels   14;  Gaps    2;

Qy         13 FVQGLEHTHRRGSLDK------RHGEGTFTSDVSSYLEGQAAKEFIAWLVKG-------- 58
              | | | :: | | : :      || |||:|:||: :|| :|||||| ||:||
Db          2 FAQWLMNSKRSGGMSRRNVQFERHAEGTYTNDVTQFLEEKAAKEFIDWLLKGIPKKQRLS 61

Qy         59 RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              || |||||||::|:|| :|||||: ||:|||
Db         62 RHAEGTFTSDMTSFLEEKAAKEFVDWLIKGR 92


RESULT 13
GLUC1_XENLA
ID   GLUC1_XENLA    STANDARD;      PRT;   266 AA.
AC   O42143;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Glucagon I precursor [Contains: Glucagon; Glucagon-like peptide 1A
DE   (GLP-1A); Glucagon-like peptide 1B (GLP-1B); Glucagon-like peptide 1C
DE   (GLP-1C); Glucagon-like peptide 2 (GLP-2)].
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
RC   TISSUE=Pancreas;
RX   MEDLINE=97368292; PubMed=9223287; DOI=10.1073/pnas.94.15.7915;
RA   Irwin D.M., Satkunarajah M., Wen Y., Brubaker P.L., Pederson R.A.,
RA   Wheeler M.B.;
RT   "The Xenopus proglucagon gene encodes novel GLP-1-like peptides with
RT   insulinotropic properties.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:7915-7920(1997).
CC   -!- FUNCTION: Promotes hydrolysis of glycogen and lipids, and raises
CC       the blood sugar level.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=O42143-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=O42143-2; Sequence=VSP_001755;
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------------
DR   EMBL; AF004432; AAB65660.1; -; mRNA.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 5.
DR   PRINTS; PR00275; GLUCAGON.
DR   PROSITE; PS00260; GLUCAGON; 5.
KW   Alternative splicing; Cleavage on pair of basic residues;
KW   Glucagon family; Hormone; Multigene family; Signal.
FT   SIGNAL       1     20       Potential.
FT   PROPEP      21     50
FT   PEPTIDE      53     81       Glucagon.
FT   PROPEP      84     95
FT   PEPTIDE      97    133       Glucagon-like peptide 1A.
FT   PROPEP     136    140
FT   PEPTIDE     142    172       Glucagon-like peptide 1B.
FT   PROPEP     175    178
FT   PEPTIDE     180    210       Glucagon-like peptide 1C.
FT   PROPEP     213    224
FT   PEPTIDE     227    259       Glucagon-like peptide 2.
FT   PROPEP     261    266
FT   VARSPLIC    214    261       Missing (in isoform 2).
FT                               /FTId=VSP_001755.
SQ   SEQUENCE   266 AA;  30951 MW;  544F7BBC20AF872C CRC64;

  Query Match             51.2%;  Score 238;  DB 1;  Length 266;
  Best Local Similarity   57.9%;  Pred. No. 8.3e-17;
  Matches   44;  Conservative   14;  Mismatches   10;  Indels    8;  Gaps    1;

Qy         22 RRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKG--------RHGEGTFTSDVSSYL 73
              ||  :  :|| |||:|:||: ||| :|||||| ||:||        || |||||:|:::||
Db        134 RRNAEFERHAEGTYTNDVTEYLEEKAAKEFIEWLIKGKPKKIRYSRHAEGTFTNDMTNYL 193

Qy         74 EGQAAKEFIAWLVKGR 89
              | :||||||: ||:|||
Db        194 EEKAAKEFVGWLIKGR 209


RESULT 14
Q6DIZ4_XENTR
ID   Q6DIZ4_XENTR PRELIMINARY;      PRT;   266 AA.
AC   Q6DIZ4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Glucagon.
GN   Name=gcg-prov;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Silurana.
OX   NCBI_TaxID=8364;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Whole body;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC075391; AAH75391.1; -; mRNA.
DR   GO; GO:0005576; C:extracellular region; IEA.
DR   GO; GO:0005179; F:hormone activity; IEA.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 5.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 5.
DR   PROSITE; PS00260; GLUCAGON; 5.
SQ   SEQUENCE   266 AA;  30809 MW;  47BAE7DD28EFF7EA CRC64;

  Query Match             50.2%;  Score 233.5;  DB 2;  Length 266;
  Best Local Similarity   48.0%;  Pred. No. 2.5e-16;
  Matches   47;  Conservative   14;  Mismatches   16;  Indels   21;  Gaps    2;

Qy         13 FVQGLEHTHRRGSLDK------RHGEGTFTSDVSSYLEGQAAKEFIAWLVKG-------- 58
              |:| | :| | | | :      || |||||||: :|: :|||||| ||: |
Db         74 FIQWLMNTKRSGGLSRRNADYERHAEGTFTSDVTQHLDEKAAKEFIDWLINGGPSKEIIS 133

Qy         59 -------RHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
                    || |||:|:||: ||| :|||||| ||: |:
Db        134 RRNADVERHAEGTYTNDVTEYLEEKAAKEFIEWLINGK 171


RESULT 15
GLUC_OCTDE
ID   GLUC_OCTDE     STANDARD;      PRT;   180 AA.
AC   P22890;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE   (GRPP); Oxyntomodulin (OXY) (OXM); Glucagon; Glucagon-like peptide 1
DE   (GLP-1); Glucagon-like peptide 1(7-37) (GLP-1(7-37)); Glucagon-like
DE   peptide 1(7-36) (GLP-1(7-36)); Glucagon-like peptide 2 (GLP-2)].
GN   Name=GCG;
OS   Octodon degus (Degu).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
OC   Hystricognathi; Octodontidae; Octodon.
OX   NCBI_TaxID=10160;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=91155952; PubMed=2293024;
RA   Nishi M., Steiner D.F.;
RT   "Cloning of complementary DNAs encoding islet amyloid polypeptide,
RT   insulin, and glucagon precursors from a New World rodent, the degu,
RT   Octodon degus.";
RL   Mol. Endocrinol. 4:1192-1198(1990).
RN   [2]
RP   REVIEW.
RX   PubMed=12554744; DOI=10.1210/me.2002-0306;
RA   Drucker D.J.;
RT   "Glucagon-like peptides: regulators of cell proliferation,
RT   differentiation, and apoptosis.";
RL   Mol. Endocrinol. 17:161-171(2003).
RN   [3]
RP   REVIEW.
RX   PubMed=12626323; DOI=10.1152/ajpendo.00492.2002;
RA   Jiang G., Zhang B.B.;
RT   "Glucagon and regulation of glucose metabolism.";
RL   Am. J. Physiol. 284:E671-E678(2003).
```

```
RN   [4]
RP   REVIEW.
RX   PubMed=10322410;
RA   Drucker D.J.;
RT   "Glucagon-like peptide 2.";
RL   Trends Endocrinol. Metab. 10:153-156(1999).
RN   [5]
RP   REVIEW.
RX   PubMed=10605628; DOI=10.1210/er.20.6.876;
RA   Kieffer T.J., Habener J.F.;
RT   "The glucagon-like peptides.";
RL   Endocr. Rev. 20:876-913(1999).
CC   -!- FUNCTION: Glucagon plays a key role in glucose metabolism and
CC       homeostasis. Regulates blood glucose by increasing gluconeogenesis
CC       and decreasing glycolysis. A counterregulatory hormone of insulin,
CC       raises plasma glucose levels in response to insulin-induced
CC       hypoglycemia (By similarity).
CC   -!- FUNCTION: GLP-1 is a potent stimulator of glucose-dependent
CC       insulin release. Play important roles on gastric motility and the
CC       suppression of plasma glucagon levels. May be involved in the
CC       suppression of satiety and stimulation of glucose disposal in
CC       peripheral tissues, independent of the actions of insulin. Have
CC       growth-promoting activities on intestinal epithelium. May also
CC       regulate the hypothalamic pituitary axis (HPA) via effects on LH,
CC       TSH, CRH, oxytocin, and vasopressin (By similarity).
CC   -!- FUNCTION: GLP-2 stimulates intestinal growth and up-regulates
CC       villus height in the small intestine, concomitant with increased
CC       crypt cell proliferation and decreased enterocyte apoptosis. The
CC       gastrointestinal tract, from the stomach to the colon is the
CC       principal target for GLP-2 action. Plays a key role in nutrient
CC       homeostasis, enhancing nutrient assimilation through enhanced
CC       gastrointestinal function, as well as increasing nutrient
CC       disposal. Stimulates intestinal glucose transport and decreases
CC       mucosal permeability (By similarity).
CC   -!- FUNCTION: Glicentin may modulate gastric acid secretion and
CC       gastro-pyloro-duodenal activity (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Glucagon is secreted in the A cells of the
CC       islets of Langerhans. GLP-1, GLP-2, oxyntomodulin and glicentin
CC       are secreted from enteroendocrine cells throughout the
CC       gastrointestinal tract. GLP1 and GLP2 are also secreted in
CC       selected neurons in the brain.
CC   -!- INDUCTION: Glucagon release is stimulated by hypoglycemia and
CC       inhibited by hyperglycemia, insulin, and somatostatin. GLP-1 and
CC       GLP-2 are induced in response to nutrient ingestion (By
CC       similarity).
CC   -!- PTM: Proglucagon is posttranslationally processed in a tissue-
CC       specific manner in pancreatic A cells and intestinal L cells. In
CC       pancreatic A cells, the major bioactive hormone is glucagon
CC       cleaved by PCSK2/PC2. In the intestinal L cells PCSK1/PC1
CC       liberates GLP-1, GLP-2, glicentin and oxyntomodulin. GLP-1 is
CC       further N-terminally truncated by posttranslational processing in
CC       the intestinal L cells resulting in GLP-1(7-37) GLP-1-(7-36)amide.
CC       The C-terminal amidation is neither important for the metabolism
CC       of GLP-1 nor for its effects on the endocrine pancreas (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; M57688; AAA40588.1; -; mRNA.
DR   PIR; C36118; GCRTDU.
DR   HSSP; P01275; 1D0R.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Amidation; Cleavage on pair of basic residues; Glucagon family;
KW   Hormone; Signal.
```

```
FT   SIGNAL       1     20
FT   PEPTIDE      21    89       Glicentin (By similarity).
FT   PEPTIDE      21    50       Glicentin-related polypeptide (By
FT                               similarity).
FT   PEPTIDE      53    89       Oxyntomodulin (By similarity).
FT   PEPTIDE      53    81       Glucagon (By similarity).
FT   PROPEP       84    89       By similarity.
FT   PEPTIDE      92    128      Glucagon-like peptide 1 (By similarity).
FT   PEPTIDE      98    128      Glucagon-like peptide 1(7-37) (By
FT                               similarity).
FT   PEPTIDE      98    127      Glucagon-like peptide 1(7-36) (By
FT                               similarity).
FT   PROPEP       131   145      By similarity.
FT   PEPTIDE      146   178      Glucagon-like peptide 2 (By similarity).
FT   SITE         52    53       Cleavage (by PCSK2) (By similarity).
FT   SITE         83    84       Cleavage (by PCSK1 and PCSK2) (By
FT                               similarity).
FT   SITE         91    92       Cleavage (by PCSK1) (By similarity).
FT   SITE         97    98       Cleavage (by PCSK1) (By similarity).
FT   SITE         130   131      Cleavage (by PCSK1) (By similarity).
FT   SITE         145   146      Cleavage (by PCSK1) (By similarity).
FT   MOD_RES      127   127      Arginine amide (G-128 provides amide
FT                               group) (By similarity).
SQ   SEQUENCE   180 AA;  21166 MW;  6E8836160A9A3051 CRC64;

  Query Match             49.6%;  Score 230.5;  DB 1;  Length 180;
  Best Local Similarity   59.0%;  Pred. No. 3.3e-16;
  Matches   46;  Conservative    7;  Mismatches   10;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSS 71
              |||| :|||||| | :|: : |::|: ||                || ||||||||||
Db         50 DKRHSQGTFTSDYSKFLDTRRAQDFLDWLKNTKRNRNNEIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


Search completed: April 19, 2006, 12:08:52
Job time : 24.5142 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:09:12 ; Search time 5.62571 Seconds
(without alignments)
1307.948 Million cell updates/sec

Title: US-10-775-180-449
Perfect score: 465
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....SSYLEGQAAKEFIAWLVKGR 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	243.5	52.4	180	2	US-08-784-582-58	Sequence 58, Appl
2	241.5	51.9	180	2	US-08-784-582-56	Sequence 56, Appl
3	241.5	51.9	180	2	US-09-635-679E-2	Sequence 2, Appli
4	237.5	51.1	180	2	US-08-784-582-61	Sequence 61, Appl
5	235.5	50.6	360	2	US-08-784-582-73	Sequence 73, Appl
6	164	35.3	36	2	US-09-614-847-119	Sequence 119, App
7	164	35.3	42	2	US-09-614-847-118	Sequence 118, App
8	157	33.8	30	2	US-09-209-799D-15	Sequence 15, Appl
9	157	33.8	30	2	US-09-614-847-87	Sequence 87, Appl
10	157	33.8	30	2	US-09-614-847-112	Sequence 112, App
11	157	33.8	30	2	US-09-614-847-113	Sequence 113, App
12	157	33.8	30	2	US-09-997-792A-13	Sequence 13, Appl
13	157	33.8	31	2	US-09-209-799D-16	Sequence 16, Appl
14	157	33.8	31	2	US-09-614-847-111	Sequence 111, App
15	157	33.8	31	2	US-09-614-847-123	Sequence 123, App
16	157	33.8	31	2	US-09-997-792A-14	Sequence 14, Appl
17	157	33.8	32	2	US-09-614-847-147	Sequence 147, App
18	157	33.8	34	2	US-09-212-663-25	Sequence 25, Appl
19	157	33.8	36	1	US-08-095-162-15	Sequence 15, Appl
20	157	33.8	36	1	US-08-470-220A-15	Sequence 15, Appl
21	157	33.8	36	1	US-08-808-825-9	Sequence 9, Appli
22	157	33.8	36	1	US-08-899-324-1	Sequence 1, Appli
23	157	33.8	36	2	US-08-967-374-15	Sequence 15, Appl
24	157	33.8	36	2	US-08-329-892B-1	Sequence 1, Appli
25	157	33.8	36	2	US-09-302-596-2	Sequence 2, Appli
26	157	33.8	36	2	US-08-472-349-6	Sequence 6, Appli
27	157	33.8	36	2	US-09-333-415-2	Sequence 2, Appli
28	157	33.8	36	2	US-09-505-991-15	Sequence 15, Appl
29	157	33.8	36	2	US-09-303-016-2	Sequence 2, Appli
30	157	33.8	36	2	US-09-614-847-88	Sequence 88, Appl
31	157	33.8	36	2	US-09-614-847-90	Sequence 90, Appl
32	157	33.8	36	2	US-09-614-847-103	Sequence 103, App
33	157	33.8	36	2	US-09-805-507-2	Sequence 2, Appli
34	157	33.8	36	2	US-09-859-804-2	Sequence 2, Appli
35	157	33.8	36	2	US-09-943-084-6	Sequence 6, Appli
36	157	33.8	36	2	US-09-623-548A-342	Sequence 342, App
37	157	33.8	36	2	US-09-623-548A-354	Sequence 354, App
38	157	33.8	36	2	US-10-055-259-2	Sequence 2, Appli
39	157	33.8	36	2	US-09-657-276-342	Sequence 342, App
40	157	33.8	36	2	US-09-657-276-354	Sequence 354, App
41	157	33.8	36	2	US-09-982-978-2	Sequence 2, Appli
42	157	33.8	36	4	PCT-US95-15800-24	Sequence 24, Appl
43	157	33.8	37	1	US-08-095-162-19	Sequence 19, Appl
44	157	33.8	37	1	US-08-470-220A-19	Sequence 19, Appl
45	157	33.8	37	1	US-08-807-263-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-784-582-58
; Sequence 58, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  180 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
```

```
;      TOPOLOGY:  linear
US-08-784-582-58

  Query Match             52.4%;  Score 243.5;  DB 2;  Length 180;
  Best Local Similarity   42.4%;  Pred. No. 2.3e-21;
  Matches   53;  Conservative   15;  Mismatches   20;  Indels   37;  Gaps    3;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRG-----------------SLDKRHGEGTFTSDV 39
              :|:::    :  |||     |: |  :                  : |||| :||||||
Db          3 SIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDPDQMNEDKRHSQGTFTSDY 62

Qy         40 SSYLEGQAAKEFIAWLVK---------------GRHGEGTFTSDVSSYLEGQAAKEFIAW 84
              | ||: : |::|: ||:                || ||||||||||||||||||||||
Db         63 SKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAW 122

Qy         85 LVKGR 89
              |||||
Db        123 LVKGR 127


RESULT 2
US-08-784-582-56
; Sequence 56, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  180 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-784-582-56

  Query Match             51.9%;  Score 241.5;  DB 2;  Length 180;
  Best Local Similarity   60.3%;  Pred. No. 4e-21;
  Matches   47;  Conservative    7;  Mismatches    9;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSS 71
              |||| :|||||| | ||: : |::|: ||:                || ||||||||||
Db         50 DKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


RESULT 3
US-09-635-679E-2
; Sequence 2, Application US/09635679E
; Patent No. 6849708
; GENERAL INFORMATION:
;  APPLICANT: Habener, Joel
;  TITLE OF INVENTION: Insulinotropic Hormone and Uses Thereof
;  FILE REFERENCE: 0609.1090009
;  CURRENT APPLICATION NUMBER: US/09/635,679E
;  CURRENT FILING DATE:  2000-08-10
;  PRIOR APPLICATION NUMBER: 09/090,949
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 08/749,762
;  PRIOR FILING DATE: 1996-11-20
;  PRIOR APPLICATION NUMBER: 08/156,800
;  PRIOR FILING DATE: 1993-11-23
;  PRIOR APPLICATION NUMBER: 07/756,215
;  PRIOR FILING DATE: 1991-09-05
;  PRIOR APPLICATION NUMBER: 07/532,111
;  PRIOR FILING DATE: 1990-06-01
;  PRIOR APPLICATION NUMBER: 07/148,517
;  PRIOR FILING DATE: 1988-01-26
;  PRIOR APPLICATION NUMBER: 06/859,928
;  PRIOR FILING DATE: 1986-05-05
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 180
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: preproglucagon precursor
US-09-635-679E-2

  Query Match             51.9%;  Score 241.5;  DB 2;  Length 180;
  Best Local Similarity   60.3%;  Pred. No. 4e-21;
  Matches   47;  Conservative    7;  Mismatches    9;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSS 71
              |||| :|||||| | ||: : |::|: ||:                || ||||||||||
Db         50 DKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSS 109

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        110 YLEGQAAKEFIAWLVKGR 127


RESULT 4
US-08-784-582-61
; Sequence 61, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
```

```
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  180 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-784-582-61

  Query Match             51.1%;  Score 237.5;  DB 2;  Length 180;
  Best Local Similarity   41.6%;  Pred. No. 1.2e-20;
  Matches   52;  Conservative   15;  Mismatches   21;  Indels   37;  Gaps    3;

Qy          2 NIFYIFLFLLSFVQG-----LEHTHRRG----------------SLDKRHGEGTFTSDV 39
              :|:::    :  |||     |: |  :                : || | :||||||
Db          3 SIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDPDQMNEDKAHSQGTFTSDY 62

Qy         40 SSYLEGQAAKEFIAWLVK--------------GRHGEGTFTSDVSSYLEGQAAKEFIAW 84
              | ||: : |::|: ||:                || ||||||||||||||||||||
Db         63 SKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSSYLEGQAAKEFIAW 122

Qy         85 LVKGR 89
              |||||
Db        123 LVKGR 127


RESULT 5
US-08-784-582-73
; Sequence 73, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  73:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  360 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-784-582-73

  Query Match             50.6%;  Score 235.5;  DB 2;  Length 360;
  Best Local Similarity   59.0%;  Pred. No. 5.1e-20;
  Matches   46;  Conservative    7;  Mismatches   10;  Indels   15;  Gaps    1;

Qy         27 DKRHGEGTFTSDVSSYLEGQAAKEFIAWLVK---------------GRHGEGTFTSDVSS 71
              || | :|||||| | ||: : |::|: ||:                 || ||||||||||
Db        230 DKAHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGTFTSDVSS 289

Qy         72 YLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||
Db        290 YLEGQAAKEFIAWLVKGR 307


RESULT 6
US-09-614-847-119
; Sequence 119, Application US/09614847
; Patent No. 6528486
; GENERAL INFORMATION:
;  APPLICANT: Larsen, Bjarne Due
;  APPLICANT:  Mikkelsen, Jens Mollgaard
;  APPLICANT:  Neve, Soren
;  TITLE OF INVENTION: NOVEL PEPTIDE AGONISTS OF GLP-1 ACTIVITY
;  FILE REFERENCE: 55511(45487)
;  CURRENT APPLICATION NUMBER: US/09/614,847
;  CURRENT FILING DATE:  2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/143,591
;  PRIOR FILING DATE: 1999-07-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
;   LENGTH: 36
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Lys6-Gly8-GLP-1(7-36)
```

US-09-614-847-119

Query Match 35.3%; Score 164; DB 2; Length 36;
Best Local Similarity 96.9%; Pred. No. 9.4e-13;
Matches 31; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy         28 KRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              |:||||||||||||||||||||||||||||||
Db          5 KKHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 36
```

RESULT 7
US-09-614-847-118
; Sequence 118, Application US/09614847
; Patent No. 6528486
; GENERAL INFORMATION:
; APPLICANT: Larsen, Bjarne Due
; APPLICANT: Mikkelsen, Jens Mollgaard
; APPLICANT: Neve, Soren
; TITLE OF INVENTION: NOVEL PEPTIDE AGONISTS OF GLP-1 ACTIVITY
; FILE REFERENCE: 55511(45487)
; CURRENT APPLICATION NUMBER: US/09/614,847
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/143,591
; PRIOR FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 118
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Lys6-Gly8-GLP-1(7-36)-Lys6
US-09-614-847-118

Query Match 35.3%; Score 164; DB 2; Length 42;
Best Local Similarity 96.9%; Pred. No. 1.1e-12;
Matches 31; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy         28 KRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              |:||||||||||||||||||||||||||||||
Db          5 KKHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 36
```

RESULT 8
US-09-209-799D-15
; Sequence 15, Application US/09209799D
; Patent No. 6380357
; GENERAL INFORMATION:
; APPLICANT: Hermeling, Ronald
; APPLICANT: Hoffmann, James
; APPLICANT: Narasimhan, Chakravarthy
; TITLE OF INVENTION: GLUCAGON-LIKE PEPTIDE-1 CRYSTALS
; FILE REFERENCE: X-10242
; CURRENT APPLICATION NUMBER: US/09/209,799D
; CURRENT FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-209-799D-15

Query Match 33.8%; Score 157; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.1e-12;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              ||||||||||||||||||||||||||||||
Db          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30
```

RESULT 9
US-09-614-847-87
; Sequence 87, Application US/09614847
; Patent No. 6528486
; GENERAL INFORMATION:
; APPLICANT: Larsen, Bjarne Due
; APPLICANT: Mikkelsen, Jens Mollgaard
; APPLICANT: Neve, Soren
; TITLE OF INVENTION: NOVEL PEPTIDE AGONISTS OF GLP-1 ACTIVITY
; FILE REFERENCE: 55511(45487)
; CURRENT APPLICATION NUMBER: US/09/614,847
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/143,591
; PRIOR FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Gly8-GLP-1-(7-36)(Human)-NH2
US-09-614-847-87

Query Match 33.8%; Score 157; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.1e-12;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              ||||||||||||||||||||||||||||||
Db          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30
```

RESULT 10
US-09-614-847-112
; Sequence 112, Application US/09614847
; Patent No. 6528486
; GENERAL INFORMATION:
; APPLICANT: Larsen, Bjarne Due
; APPLICANT: Mikkelsen, Jens Mollgaard
; APPLICANT: Neve, Soren
; TITLE OF INVENTION: NOVEL PEPTIDE AGONISTS OF GLP-1 ACTIVITY
; FILE REFERENCE: 55511(45487)
; CURRENT APPLICATION NUMBER: US/09/614,847
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/143,591
; PRIOR FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 112
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Gly8Lys34N-palmitoyl-GLP-1 (7-36)
; NAME/KEY: MOD_RES
; LOCATION: (28)
; OTHER INFORMATION: Lys(N-palmitoyl)
US-09-614-847-112

Query Match 33.8%; Score 157; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.1e-12;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              ||||||||||||||||||||||||||||||
Db          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30
```

```
RESULT 11
US-09-614-847-113
; Sequence 113, Application US/09614847
; Patent No. 6528486
; GENERAL INFORMATION:
;  APPLICANT: Larsen, Bjarne Due
;  APPLICANT:  Mikkelsen, Jens Mollgaard
;  APPLICANT:  Neve, Soren
;  TITLE OF INVENTION: NOVEL PEPTIDE AGONISTS OF GLP-1 ACTIVITY
;  FILE REFERENCE: 55511(45487)
;  CURRENT APPLICATION NUMBER: US/09/614,847
;  CURRENT FILING DATE:  2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/143,591
;  PRIOR FILING DATE: 1999-07-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Gly8Lys26N-palmitoyl-GLP-1 (7-36)
;   NAME/KEY: MOD_RES
;   LOCATION: (20)
;   OTHER INFORMATION: Lys(N-palmitoyl)
US-09-614-847-113

  Query Match             33.8%;  Score 157;  DB 2;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-12;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              ||||||||||||||||||||||||||||||
Db          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 12
US-09-997-792A-13
; Sequence 13, Application US/09997792A
; Patent No. 6555521
; GENERAL INFORMATION:
;  APPLICANT: ELI LILLY and COMPANY
;  TITLE OF INVENTION: Glucagon-Like Peptide-1 Crystals
;  FILE REFERENCE: X-10242A
;  CURRENT APPLICATION NUMBER: US/09/997,792A
;  CURRENT FILING DATE:  2002-09-30
;  PRIOR APPLICATION NUMBER: US 60/069,728
;  PRIOR FILING DATE: 1997-12-16
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Construct
US-09-997-792A-13

  Query Match             33.8%;  Score 157;  DB 2;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-12;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              ||||||||||||||||||||||||||||||
Db          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 13
US-09-209-799D-16
; Sequence 16, Application US/09209799D
; Patent No. 6380357
; GENERAL INFORMATION:
;  APPLICANT: Hermeling, Ronald
;  APPLICANT:  Hoffmann, James
;  APPLICANT:  Narasimhan, Chakravarthy
;  TITLE OF INVENTION: GLUCAGON-LIKE PEPTIDE-1 CRYSTALS
;  FILE REFERENCE: X-10242
;  CURRENT APPLICATION NUMBER: US/09/209,799D
;  CURRENT FILING DATE:  1998-12-11
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: synthetic construct
US-09-209-799D-16

  Query Match             33.8%;  Score 157;  DB 2;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-12;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              ||||||||||||||||||||||||||||||
Db          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 14
US-09-614-847-111
; Sequence 111, Application US/09614847
; Patent No. 6528486
; GENERAL INFORMATION:
;  APPLICANT: Larsen, Bjarne Due
;  APPLICANT:  Mikkelsen, Jens Mollgaard
;  APPLICANT:  Neve, Soren
;  TITLE OF INVENTION: NOVEL PEPTIDE AGONISTS OF GLP-1 ACTIVITY
;  FILE REFERENCE: 55511(45487)
;  CURRENT APPLICATION NUMBER: US/09/614,847
;  CURRENT FILING DATE:  2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/143,591
;  PRIOR FILING DATE: 1999-07-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 111
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Gly8Lys37N-palmitoyl-GLP-1 (7-36)
;   NAME/KEY: MOD_RES
;   LOCATION: (31)
;   OTHER INFORMATION: Lys(N-palmitoyl)
US-09-614-847-111

  Query Match             33.8%;  Score 157;  DB 2;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-12;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              ||||||||||||||||||||||||||||||
Db          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 15
US-09-614-847-123
; Sequence 123, Application US/09614847
; Patent No. 6528486
; GENERAL INFORMATION:
;  APPLICANT: Larsen, Bjarne Due
;  APPLICANT:  Mikkelsen, Jens Mollgaard
;  APPLICANT:  Neve, Soren
;  TITLE OF INVENTION: NOVEL PEPTIDE AGONISTS OF GLP-1 ACTIVITY
```

```
;  FILE REFERENCE: 55511(45487)
;  CURRENT APPLICATION NUMBER: US/09/614,847
;  CURRENT FILING DATE:  2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/143,591
;  PRIOR FILING DATE: 1999-07-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 123
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Gly8-GLP-1(7-37)
US-09-614-847-123

  Query Match             33.8%;  Score 157;  DB 2;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-12;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              ||||||||||||||||||||||||||||||
Db          1 HGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


Search completed: April 19, 2006, 12:11:37
Job time : 6.62571 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:29:13 ; Search time 18.7734 Seconds
(without alignments)
1980.821 Million cell updates/sec

Title: US-10-775-180-449
Perfect score: 465
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....SSYLEGQAAKEFIAWLVKGR 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	465	100.0	89	5	US-10-775-180-449	Sequence 449, App
2	465	100.0	89	5	US-10-775-204-1282	Sequence 1282, Ap
3	465	100.0	673	5	US-10-775-180-834	Sequence 834, App
4	465	100.0	673	5	US-10-775-204-2170	Sequence 2170, Ap
5	465	100.0	674	5	US-10-775-180-447	Sequence 447, App
6	465	100.0	674	5	US-10-775-204-1280	Sequence 1280, Ap
7	341.5	73.4	669	5	US-10-775-180-419	Sequence 419, App
8	341.5	73.4	669	5	US-10-775-204-1231	Sequence 1231, Ap
9	341	73.3	145	5	US-10-775-180-685	Sequence 685, App
10	341	73.3	145	5	US-10-775-204-1790	Sequence 1790, Ap
11	341	73.3	730	5	US-10-775-180-610	Sequence 610, App
12	341	73.3	730	5	US-10-775-204-1622	Sequence 1622, Ap
13	335.5	72.2	669	5	US-10-775-180-425	Sequence 425, App
14	335.5	72.2	669	5	US-10-775-204-1237	Sequence 1237, Ap
15	335	72.0	145	5	US-10-775-180-687	Sequence 687, App
16	335	72.0	145	5	US-10-775-204-1792	Sequence 1792, Ap
17	335	72.0	730	5	US-10-775-180-612	Sequence 612, App
18	335	72.0	730	5	US-10-775-204-1624	Sequence 1624, Ap
19	329.5	70.9	669	5	US-10-775-180-420	Sequence 420, App
20	329.5	70.9	669	5	US-10-775-180-421	Sequence 421, App
21	329.5	70.9	669	5	US-10-775-180-423	Sequence 423, App
22	329.5	70.9	669	5	US-10-775-180-424	Sequence 424, App
23	329.5	70.9	669	5	US-10-775-204-1232	Sequence 1232, Ap
24	329.5	70.9	669	5	US-10-775-204-1233	Sequence 1233, Ap
25	329.5	70.9	669	5	US-10-775-204-1235	Sequence 1235, Ap
26	329.5	70.9	669	5	US-10-775-204-1236	Sequence 1236, Ap
27	324	69.7	83	5	US-10-775-180-684	Sequence 684, App
28	324	69.7	83	5	US-10-775-204-1789	Sequence 1789, Ap
29	324	69.7	668	5	US-10-775-180-609	Sequence 609, App
30	324	69.7	668	5	US-10-775-204-1621	Sequence 1621, Ap
31	319	68.6	77	5	US-10-775-180-686	Sequence 686, App
32	319	68.6	77	5	US-10-775-204-1791	Sequence 1791, Ap
33	319	68.6	662	5	US-10-775-180-611	Sequence 611, App
34	319	68.6	662	5	US-10-775-204-1623	Sequence 1623, Ap
35	318	68.4	83	5	US-10-775-180-688	Sequence 688, App
36	318	68.4	83	5	US-10-775-204-1793	Sequence 1793, Ap
37	318	68.4	668	5	US-10-775-180-613	Sequence 613, App
38	318	68.4	668	5	US-10-775-204-1625	Sequence 1625, Ap
39	317.5	68.3	664	5	US-10-775-180-598	Sequence 598, App
40	317.5	68.3	664	5	US-10-775-204-1607	Sequence 1607, Ap
41	315.5	67.8	663	5	US-10-775-180-600	Sequence 600, App
42	315.5	67.8	663	5	US-10-775-204-1609	Sequence 1609, Ap
43	314	67.5	60	5	US-10-775-180-835	Sequence 835, App
44	314	67.5	60	5	US-10-775-204-2180	Sequence 2180, Ap
45	313	67.3	77	5	US-10-775-180-689	Sequence 689, App

ALIGNMENTS

RESULT 1
US-10-775-180-449
; Sequence 449, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 449
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-180-449

Query Match 100.0%; Score 465; DB 5; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.8e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
```

```
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 2
US-10-775-204-1282
; Sequence 1282, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1282
;   LENGTH: 89
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1282

  Query Match             100.0%;  Score 465;  DB 5;  Length 89;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-44;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 3
US-10-775-180-834
; Sequence 834, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 834
;   LENGTH: 673
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-834

  Query Match             100.0%;  Score 465;  DB 5;  Length 673;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-43;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 4
US-10-775-204-2170
; Sequence 2170, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
```

```
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2170
;   LENGTH: 673
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-2170

  Query Match             100.0%;  Score 465;  DB 5;  Length 673;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-43;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 5
US-10-775-180-447
; Sequence 447, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 447
;   LENGTH: 674
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-447

  Query Match             100.0%;  Score 465;  DB 5;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-43;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 6
US-10-775-204-1280
; Sequence 1280, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1280
;   LENGTH: 674
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1280

  Query Match             100.0%;  Score 465;  DB 5;  Length 674;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-43;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 7
US-10-775-180-419
; Sequence 419, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
```

```
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 419
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-419

  Query Match             73.4%;  Score 341.5;  DB 5;  Length 669;
  Best Local Similarity   79.3%;  Pred. No. 1.9e-29;
  Matches   69;  Conservative    4;  Mismatches    5;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||||||||||||||||||||||||||||||||||
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGR 84


RESULT 8
US-10-775-204-1231
; Sequence 1231, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1231
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1231

  Query Match             73.4%;  Score 341.5;  DB 5;  Length 669;
  Best Local Similarity   79.3%;  Pred. No. 1.9e-29;
  Matches   69;  Conservative    4;  Mismatches    5;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         ||||||||||||||||||||||||||||||||||||
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGR 84


RESULT 9
US-10-775-180-685
; Sequence 685, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 685
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-685

  Query Match             73.3%;  Score 341;  DB 5;  Length 145;
  Best Local Similarity   81.7%;  Pred. No. 3.7e-30;
  Matches   67;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :        ||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGR 145


RESULT 10
US-10-775-204-1790
; Sequence 1790, Application US/10775204
```

```
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1790
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1790

  Query Match             73.3%;  Score 341;  DB 5;  Length 145;
  Best Local Similarity   81.7%;  Pred. No. 3.7e-30;
  Matches   67;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGR 145


RESULT 11
US-10-775-180-610
; Sequence 610, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 610
;   LENGTH: 730
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-610

  Query Match             73.3%;  Score 341;  DB 5;  Length 730;
  Best Local Similarity   81.7%;  Pred. No. 2.4e-29;
  Matches   67;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGR 145


RESULT 12
US-10-775-204-1622
; Sequence 1622, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1622
;   LENGTH: 730
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1622
```

```
  Query Match             73.3%;  Score 341;  DB 5;  Length 730;
  Best Local Similarity   81.7%;  Pred. No. 2.4e-29;
  Matches   67;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       ||||||||||||||||||||||||||||||||||||||||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGR 145


RESULT 13
US-10-775-180-425
; Sequence 425, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 425
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-425

  Query Match             72.2%;  Score 335.5;  DB 5;  Length 669;
  Best Local Similarity   78.2%;  Pred. No. 8.8e-29;
  Matches   68;  Conservative    4;  Mismatches    6;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         |||||||||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGR 84


RESULT 14
US-10-775-204-1237
; Sequence 1237, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Balance, David J.
;  APPLICANT:  Turner, Andrew J.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF564
;  CURRENT APPLICATION NUMBER: US/10/775,204
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: 60/351,360
;  PRIOR FILING DATE: 2002-01-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2222
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1237
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-204-1237

  Query Match             72.2%;  Score 335.5;  DB 5;  Length 669;
  Best Local Similarity   78.2%;  Pred. No. 8.8e-29;
  Matches   68;  Conservative    4;  Mismatches    6;  Indels    9;  Gaps    1;

Qy          3 IFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGE 62
              |  :|||  :: :         |||||||||||||||||||||||||||||||||| |
Db          7 ISLLFLFSSAYSR---------SLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAE 57

Qy         63 GTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||||||
Db         58 GTFTSDVSSYLEGQAAKEFIAWLVKGR 84


RESULT 15
US-10-775-180-687
; Sequence 687, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Haseltine, William A.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF574
;  CURRENT APPLICATION NUMBER: US/10/775,180
;  CURRENT FILING DATE:  2004-02-11
;  PRIOR APPLICATION NUMBER: PCT/US02/40892
;  PRIOR FILING DATE: 2002-12-23
;  PRIOR APPLICATION NUMBER: 60/341,811
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/360,000
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: 60/378,950
;  PRIOR FILING DATE: 2002-05-10
;  PRIOR APPLICATION NUMBER: 60/398,008
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/411,355
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: 60/414,984
```

```
;  PRIOR FILING DATE: 2002-10-02
;  PRIOR APPLICATION NUMBER: 60/417,611
;  PRIOR FILING DATE: 2002-10-11
;  PRIOR APPLICATION NUMBER: 60/420,246
;  PRIOR FILING DATE: 2002-10-23
;  PRIOR APPLICATION NUMBER: 60/423,623
;  PRIOR FILING DATE: 2002-11-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 858
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 687
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-775-180-687

  Query Match             72.0%;  Score 335;  DB 5;  Length 145;
  Best Local Similarity   80.5%;  Pred. No. 1.7e-29;
  Matches   66;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          8 LFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHGEGTFTS 67
              ||: : :  :       |||||||||||||||||||||||||||||||||||| ||||||
Db         64 LFINTTIASIAAKEEGVSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAEGTFTS 123

Qy         68 DVSSYLEGQAAKEFIAWLVKGR 89
              ||||||||||||||||||||||
Db        124 DVSSYLEGQAAKEFIAWLVKGR 145


Search completed: April 19, 2006, 12:35:50
Job time : 18.7734 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 19, 2006, 12:31:03 ; Search time 2.97088 Seconds
(without alignments)
1318.215 Million cell updates/sec

Title: US-10-775-180-449
Perfect score: 465
Sequence: 1 MNIFYIFLFLLSFVQGLEHT.....SSYLEGQAAKEFIAWLVKGR 89

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225428 seqs, 44002918 residues

Total number of hits satisfying chosen parameters: 225428

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /SIDS5/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /SIDS5/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /SIDS5/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /SIDS5/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /SIDS5/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /SIDS5/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /SIDS5/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /SIDS5/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	465	100.0	674	7	US-11-175-690-206	Sequence 206, App
2	465	100.0	915	7	US-11-175-690-208	Sequence 208, App
3	317	68.2	647	7	US-11-175-690-212	Sequence 212, App
4	313.5	67.4	657	7	US-11-175-690-216	Sequence 216, App
5	313.5	67.4	657	7	US-11-175-690-303	Sequence 303, App
6	313	67.3	649	7	US-11-175-690-213	Sequence 213, App
7	312.5	67.2	652	7	US-11-175-690-218	Sequence 218, App
8	310.5	66.8	650	7	US-11-175-690-209	Sequence 209, App
9	309.5	66.6	646	7	US-11-175-690-223	Sequence 223, App
10	309.5	66.6	659	7	US-11-175-690-221	Sequence 221, App
11	309	66.5	651	7	US-11-175-690-224	Sequence 224, App
12	308	66.2	648	7	US-11-175-690-214	Sequence 214, App
13	308	66.2	653	7	US-11-175-690-215	Sequence 215, App
14	308	66.2	654	7	US-11-175-690-219	Sequence 219, App
15	308	66.2	655	7	US-11-175-690-220	Sequence 220, App
16	308	66.2	656	7	US-11-175-690-225	Sequence 225, App
17	308	66.2	658	7	US-11-175-690-210	Sequence 210, App
18	302.5	65.1	122	6	US-10-997-061-31	Sequence 31, Appl
19	302.5	65.1	123	6	US-10-997-074-31	Sequence 31, Appl
20	295	63.4	118	6	US-10-997-074-52	Sequence 52, Appl
21	294.5	63.3	70	6	US-10-997-061-28	Sequence 28, Appl
22	293	63.0	117	6	US-10-997-061-9	Sequence 9, Appli
23	293	63.0	117	6	US-10-997-074-28	Sequence 28, Appl
24	291	62.6	277	6	US-10-997-061-11	Sequence 11, Appl
25	290.5	62.5	287	6	US-10-997-074-55	Sequence 55, Appl
26	288	61.9	119	6	US-10-997-061-13	Sequence 13, Appl
27	243.5	52.4	180	7	US-11-145-463-1	Sequence 1, Appli
28	164	35.3	36	6	US-10-517-563-8	Sequence 8, Appli
29	164	35.3	42	6	US-10-517-563-7	Sequence 7, Appli
30	161	34.6	36	7	US-11-293-676-8	Sequence 8, Appli
31	161	34.6	39	7	US-11-293-676-9	Sequence 9, Appli
32	157	33.8	30	7	US-11-175-690-293	Sequence 293, App
33	157	33.8	30	7	US-11-175-690-295	Sequence 295, App
34	157	33.8	30	7	US-11-175-690-296	Sequence 296, App
35	157	33.8	30	7	US-11-175-690-297	Sequence 297, App
36	157	33.8	30	7	US-11-175-690-299	Sequence 299, App
37	157	33.8	30	7	US-11-175-690-300	Sequence 300, App
38	157	33.8	30	7	US-11-175-690-301	Sequence 301, App
39	157	33.8	30	7	US-11-175-690-302	Sequence 302, App
40	157	33.8	30	7	US-11-175-690-305	Sequence 305, App
41	157	33.8	30	7	US-11-175-690-306	Sequence 306, App
42	157	33.8	30	7	US-11-175-690-307	Sequence 307, App
43	157	33.8	30	7	US-11-175-690-308	Sequence 308, App
44	157	33.8	30	7	US-11-175-690-310	Sequence 310, App
45	157	33.8	30	7	US-11-175-690-311	Sequence 311, App

ALIGNMENTS

RESULT 1
US-11-175-690-206
; Sequence 206, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 206
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-206

Query Match 100.0%; Score 465; DB 7; Length 674;
Best Local Similarity 100.0%; Pred. No. 7.5e-45;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
```

RESULT 2

```
US-11-175-690-208
; Sequence 208, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 208
;   LENGTH: 915
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-208

  Query Match             100.0%;  Score 465;  DB 7;  Length 915;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-44;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89
              |||||||||||||||||||||||||||||
Db         61 GEGTFTSDVSSYLEGQAAKEFIAWLVKGR 89


RESULT 3
US-11-175-690-212
; Sequence 212, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 212
;   LENGTH: 647
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-212

  Query Match             68.2%;  Score 317;  DB 7;  Length 647;
  Best Local Similarity   75.9%;  Pred. No. 4.4e-28;
  Matches   63;  Conservative    4;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIA 83
              |:|:    :    : | |
Db         61 AHDAHKSEVAHRFKDLGEENFKA 83


RESULT 4
US-11-175-690-216
; Sequence 216, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 216
;   LENGTH: 657
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-216

  Query Match             67.4%;  Score 313.5;  DB 7;  Length 657;
  Best Local Similarity   77.1%;  Pred. No. 1.1e-27;
  Matches   64;  Conservative    4;  Mismatches   12;  Indels    3;  Gaps    1;

Qy          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIA 83
              |:|:    :    | |  :|
Db         61 AH---KSEVAHRFKDDAHKSEVA 80


RESULT 5
US-11-175-690-303
; Sequence 303, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
```

```
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 303
;   LENGTH: 657
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-303

  Query Match             67.4%;  Score 313.5;  DB 7;  Length 657;
  Best Local Similarity   77.1%;  Pred. No. 1.1e-27;
  Matches   64;  Conservative    4;  Mismatches   12;  Indels    3;  Gaps    1;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIA 83
                 |:|:   :  | |  :|
Db         61 AH---KSEVAHRFKDDAHKSEVA 80


RESULT 6
US-11-175-690-213
; Sequence 213, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 213
;   LENGTH: 649
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-213

  Query Match             67.3%;  Score 313;  DB 7;  Length 649;
  Best Local Similarity   75.3%;  Pred. No. 1.3e-27;
  Matches   64;  Conservative    4;  Mismatches   15;  Indels    2;  Gaps    1;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         60 -HGEGTFTSDVSSYLEGQAAKEFIA 83
               |      |:|:   :    : | |
Db         61 AHKSDAHKSEVAHRFKDLGEENFKA 85


RESULT 7
US-11-175-690-218
; Sequence 218, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 218
;   LENGTH: 652
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-218

  Query Match             67.2%;  Score 312.5;  DB 7;  Length 652;
  Best Local Similarity   77.8%;  Pred. No. 1.4e-27;
  Matches   63;  Conservative    4;  Mismatches   11;  Indels    3;  Gaps    1;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEF 81
                 |:|:   : : |  |
Db         61 AH---KSEVADAHKSEVAHRF 78


RESULT 8
US-11-175-690-209
; Sequence 209, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
```

```
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 209
;   LENGTH: 650
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-209

  Query Match             66.8%;  Score 310.5;  DB 7;  Length 650;
  Best Local Similarity   74.4%;  Pred. No. 2.4e-27;
  Matches   64;  Conservative    4;  Mismatches   15;  Indels    3;  Gaps    1;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR- 59
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         60 --HGEGTFTSDVSSYLEGQAAKEFIA 83
                |    |:|:   :    : | |
Db         61 AHKSEDAHKSEVAHRFKDLGEENFKA 86


RESULT 9
US-11-175-690-223
; Sequence 223, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 223
;   LENGTH: 646
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-223

  Query Match             66.6%;  Score 309.5;  DB 7;  Length 646;
  Best Local Similarity   75.9%;  Pred. No. 3.1e-27;
  Matches   63;  Conservative    5;  Mismatches   14;  Indels    1;  Gaps    1;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEFIA 83
                :    |:|:   :    : | |
Db         61 AD-AHKSEVAHRFKDLGEENFKA 82


RESULT 10
US-11-175-690-221
; Sequence 221, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 221
;   LENGTH: 659
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-175-690-221

  Query Match             66.6%;  Score 309.5;  DB 7;  Length 659;
  Best Local Similarity   76.5%;  Pred. No. 3.2e-27;
  Matches   65;  Conservative    4;  Mismatches   11;  Indels    5;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLE--GQAAKEFIA 83
                  |:|:   :  | | |  :|
Db         61 AH---KSEVAHRFKDLGDAHKSEVA 82


RESULT 11
US-11-175-690-224
; Sequence 224, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
```

```
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 224
;    LENGTH: 651
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-11-175-690-224

  Query Match             66.5%;  Score 309;  DB 7;  Length 651;
  Best Local Similarity   77.8%;  Pred. No. 3.6e-27;
  Matches   63;  Conservative    5;  Mismatches    9;  Indels    4;  Gaps    2;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGRD 60

Qy         61 GEGTFTSDVSSYLEGQAAKEF 81
                    |:| :: : : |  |
Db         61 AH---KSEVDAH-KSEVAHRF 77


RESULT 12
US-11-175-690-214
; Sequence 214, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 214
;    LENGTH: 648
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-11-175-690-214

  Query Match             66.2%;  Score 308;  DB 7;  Length 648;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-27;
  Matches   59;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59


RESULT 13
US-11-175-690-215
; Sequence 215, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 215
;    LENGTH: 653
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-11-175-690-215

  Query Match             66.2%;  Score 308;  DB 7;  Length 653;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-27;
  Matches   59;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59


RESULT 14
US-11-175-690-219
; Sequence 219, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
;  APPLICANT: Haseltine et al.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF605
;  CURRENT APPLICATION NUMBER: US/11/175,690
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: PCT/US04/001369
;  PRIOR FILING DATE: 2004-01-20
;  PRIOR APPLICATION NUMBER: US 60/441,305
;  PRIOR FILING DATE: 2003-01-22
;  PRIOR APPLICATION NUMBER: US 60/453,201
;  PRIOR FILING DATE: 2003-03-11
;  PRIOR APPLICATION NUMBER: US 60/467,222
;  PRIOR FILING DATE: 2003-05-02
;  PRIOR APPLICATION NUMBER: US 60/472,816
;  PRIOR FILING DATE: 2003-05-23
;  PRIOR APPLICATION NUMBER: US 60/476,267
;  PRIOR FILING DATE: 2003-06-06
;  PRIOR APPLICATION NUMBER: US 60/505,172
;  PRIOR FILING DATE: 2003-09-24
;  PRIOR APPLICATION NUMBER: US 60/506,746
;  PRIOR FILING DATE: 2003-09-30
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 219
```

; LENGTH: 654
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-219

Query Match 66.2%; Score 308; DB 7; Length 654;
Best Local Similarity 100.0%; Pred. No. 4.7e-27;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59

RESULT 15
US-11-175-690-220
; Sequence 220, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 220
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-220

Query Match 66.2%; Score 308; DB 7; Length 655;
Best Local Similarity 100.0%; Pred. No. 4.7e-27;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MNIFYIFLFLLSFVQGLEHTHRRGSLDKRHGEGTFTSDVSSYLEGQAAKEFIAWLVKGR 59

Search completed: April 19, 2006, 12:36:43
Job time : 3.97088 secs